Available online at www.sciencedirect.com

ScienceDirect

Mathematics and Computers in Simulation 159 (2019) 93–118

MATHEMATICS AND COMPUTERS IN SIMULATION

www.elsevier.com/locate/matcom

Original articles

M-PCM-OFFD: An effective output statistics estimation method for systems of high dimensional uncertainties subject to low-order parameter interactions

Junfei Xie[a], Yan Wan[b,*], Kevin Mills[c], James J. Filliben[c], Yu Lei[d], Zongli Lin[e]

[a] *Department of Computing Sciences, Texas A&M University-Corpus Christi, 6300 Ocean Dr, Corpus Christi, TX, 78412, United States*
[b] *Department of Electrical Engineering, University of Texas at Arlington, 701 S Nedderman Dr, Arlington, TX, 76019, United States*
[c] *National Institute of Standards and Technology, 100 Bureau Dr, Geithersburg, MD, 20899, United States*
[d] *Department of Computer Science and Engineering, University of Texas at Arlington, 701 S Nedderman Dr, Arlington, TX, 76019, United States*
[e] *Charles L. Brown Department of Electrical and Computer Engineering, University of Virginia, Charlottesville, VA, 22904, United States*

Received 17 December 2017; received in revised form 6 September 2018; accepted 27 October 2018
Available online 9 November 2018

Abstract

The evaluation of output performance statistics for systems of high-dimensional uncertain input parameters is crucial for robust real-time decision-making tasks of large-scale complex systems that operate in an uncertain environment. We develop a framework that integrates Multivariate Probabilistic Collocation Method (M-PCM) and Orthogonal Fractional Factorial Design (OFFD) to achieve an effective and scalable output statistics estimation. In this paper, we prove that when the degree of each uncertain parameter does not exceed 3 and under the widely held assumption for high-dimensional systems that the interactions among uncertain input parameters are negligible beyond certain order, the integrated M-PCM–OFFD method breaks the curse of dimensionality for correct output mean estimation by maximally reducing the number of simulations from 2^{2m} to $2^{\lceil \log_2(m+1) \rceil}$ for a system mapping of m uncertain input parameters. In addition, the resulting reduced-size simulation set is the most robust to numerical truncation errors of simulators among all subsets of the same size in the M-PCM simulation set. The analysis also provides new insightful formal interpretations of the optimality of OFFDs.
© 2018 International Association for Mathematics and Computers in Simulation (IMACS). Published by Elsevier B.V. All rights reserved.

Keywords: Output statistics estimation; Curse of dimensionality; Large-scale system simulation; Computational efficiency

1. Introduction

Large-scale complex systems (such as complex information systems, power grids, and traffic networks) commonly involve high-dimensional uncertain input parameters, which modulate system dynamics. The importance of considering these uncertainties in achieving robust decision-making solutions is increasingly recognized, when developing

* Corresponding author.
E-mail address: yan.wan@uta.edu (Y. Wan).

https://doi.org/10.1016/j.matcom.2018.10.010
0378-4754/© 2018 International Association for Mathematics and Computers in Simulation (IMACS). Published by Elsevier B.V. All rights reserved.

both model-based and data-driven learning-based decision solutions [36,66,67]. In order to achieve autonomous real-time decision-making for large-scale complex systems under uncertainty, a critical step is to develop an effective and scalable systematic procedure to evaluate statistical system performance in the presence of high-dimensional uncertain input parameters. With the real-time statistical system performance evaluation capabilities, automatic closed-loop control solutions can be made to shape the dynamics to meet desired performance requirements.

This statistical system performance evaluation problem can be formulated as the estimation of output statistics for a system of multiple uncertain input parameters with known distributions, and is typically solved using the Monte Carlo (MC) simulation method [28,55]. However, as MC requires a very large number of simulations to converge to meaningful performance estimates, it does not meet the real-time requirement, especially for large-scale complex system applications, which typically consume considerable computational time for each simulation run. Additional advanced random sampling approaches extended from MC have been developed to improve computational efficiency, including Latin Hypercube sampling [41], importance sampling [19], multilevel MC [18], greedy and adaptive sampling [24,30]. However, they still do not produce effective output-statistics estimation, that is fast in time, accurate in estimation performance, robust to numerical issues, and scalable with respect to the dimension of the uncertain input parameters.

We propose a new framework to quickly and accurately estimate the output statistics for systems of high-dimensional uncertainties. The framework, referred to as M-PCM–OFFD, integrates Multivariate Probabilistic Collocation Method (M-PCM) [70,71] with Orthogonal Fractional Factorial Designs (OFFDs) [4,22,48,49] to achieve high effectiveness. M-PCM selects a significantly reduced number of simulations compared to MC, when the dimension of uncertain input parameters is low, to construct a low-order mapping, which estimates the output mean of an original system mapping (also called functional model or response surface [5,34]) based on statistical information of uncertain input parameters. However, the number of M-PCM simulations increases exponentially as the number of uncertain input parameters increases, causing computational time issues for large-scale complex system applications. We show that by integrating M-PCM with the procedures of the experimental design method OFFD, the number of simulations can be further reduced with the accurate output mean estimation property retained, which significantly improves the scalability of M-PCM for high-dimensional uncertainty evaluation.

To the best of our knowledge, this is the first study that leverages OFFD to systematically address the scalability issue of M-PCM. Considering that the simulation of large-scale systems is often computationally intensive, reducing the number of simulation runs is a critical step towards achieving real-time output statistics estimation. In this paper, we focus on output statistics estimation for a system of m uncertain input parameters when the degree of each parameter does not exceed 3 and under the widely held assumption for high-dimensional systems that the interactions among uncertain input parameters are negligible beyond certain order. The M-PCM–OFFD framework can also be used to estimate higher output moments, and to analyze systems of higher parameter degrees and even nonlinear systems (see e.g., [65,66,71]). We discuss these aspects conceptually in this paper, and leave the complete analysis for more general systems for future work. For the system of interest in this paper, we show that the integrated M-PCM–OFFD method reduces the number of simulations required to estimate the correct output mean from 2^{2m} to the range of $\left[2^{\lceil \log_2(m+1)\rceil}, 2^{m-1}\right]$, where $\lceil x \rceil$ denotes the nearest integer above the number x. In addition, many simulators have *resolution* constraints on input parameters, and numerically truncate input parameter values to the allowed resolution levels. Such truncation may unfortunately fail the output mean estimation. We prove that the reduced-size simulation set selected by the integrated M-PCM–OFFD method is the most robust to such numerical truncation errors of simulators among all subsets of the same size in the M-PCM simulation set.

Absent results addressing the curse of dimensionality specifically for M-PCM, an output statistics estimation method, we review relevant studies in the broader directions of uncertainty evaluation and multivariate dimension reduction. Similar to M-PCM, polynomial chaos expansion (PCE) also uses the quadrature-based rules, however PCE aims to approximate the original system mapping with another mapping constructed from quadrature-based generalized polynomials and quantify uncertainty of the output [8,32,58,59,63,69]. M-PCM, on the other hand, aims for precise estimation of mean based on arbitrary probability distribution functions of uncertain input parameters, through its specific polynomial construction and truncation rules that result in a low-order mapping. With respect to PCE, when the number of Gaussian quadrature points (roots of the next higher-order orthogonal polynomials) used in these approaches is larger than the number of parameters in truncated PCE mappings (not necessarily of a lower-order as M-PCM does), several methods were introduced to reduce the number of simulations required to estimate mapping coefficients. Papers [58,59] use a subset of the Gaussian quadrature points (or collocation points) to estimate the coefficients of PCE, but the selection criteria were not specified. To improve the robustness of estimation results, paper [32] suggested a regression method, called stochastic response surface method (SRSM), that follows a heuristic procedure to select the set of simulation points that favor those in regions of high probability and of the size typically twice the number of PCE mapping coefficients. Also of relevance, when all cross-terms (i.e., terms that involve multiple parameters) are negligible, the univariate dimension reduction method (UDR) uses multiple univariate functions to approximate the original system mapping. UDR requires $mn + 1$ simulations to estimate an m-parameter system mapping, if n points are selected for each univariate function [37]. The high-dimensional model representation (HDMR) based on the analysis of variance (ANOVA) [39,50,51,54] is a more general model approximation method. It decomposes a multivariate function into a finite number of terms of increasing dimensions, where each term is a nonlinear function. By viewing the coefficient estimation problem as least-squares estimation, recent results also include the coherence-optimal sampling and its variants [14,25–27,47], a Markov Chain MC based approach that chooses a small set of samples to achieve a statistical optimality in terms of the spectral radius of the matrix constructed from samples.

By observing the dimension problem caused by the tensor product of Gaussian quadrature points, methods like monomial cubature rules (MCR), sparse grids (SG) and its extensions from the numerical integration literature were recently developed to break the curse of dimensionality [10,11,20,23,29,40,44,46,63]. In paper [63], the full set of points generated by MCR of the degree $2n + 1$ formula is used to estimate the coefficients of PCE mapping of order n through regression. However, the method is not justified. It does not produce an accurate estimation of output mean for PCEs of orders 2 and 3, and its performance for PCEs of order higher than 3 is unknown. Recently, paper [40] used SG points to calculate, one by one, the coefficients of the associated orthonormal polynomial in PCE. As each coefficient can be represented by an integral, the SG is naturally applied to approximate this integration operation. The required number of simulations depends on the accuracy level of the applied SG, and may still be large. The use of SG for pseudospectral approximation and partial differential equations are explored in [10,11,23,44]. Another related line of work is tensor decomposition [3,15,20,21,45,52], which decomposes the mapping of high dimensional inputs into mappings of low-dimensional inputs.

Our proposed M-PCM–OFFD, different from the above studies on function approximation, aims for correct output statistics estimation based on the distributions of uncertain input parameters. Based on quadrature rules, M-PCM takes advantage of the low-order mapping which only approximates the original mapping over the likely range of uncertainty parameters to estimate the output mean correctly. The "balanced" and "orthogonal" structure of OFFD further addresses the scalability issue. M-PCM–OFFD offers the following features: (1) the systematic design procedure facilitates automatic real-time performance evaluation and management under high-dimensional uncertainties, (2) arbitrary probabilistic distribution knowledge of the uncertain input parameters are allowed, (3) output mean can be precisely estimated with rigorous analysis provided, and (4) robustness to numerical truncation errors of simulators

is guaranteed. In the simulation studies, we also compare in detail the performance of M-PCM–OFFD with seven representative function approximation methods in terms of computation and estimation correctness. While we were developing this paper based on earlier results in the brief conference paper [68], we noticed a very recent development that through simulation studies show the use of OFFD to reduce the computation for PCE approximations [62], which provides further evidence for our study.

The remainder of this paper is organized as follows. In Section 2, we review fundamentals of M-PCM and OFFD. In Section 3, we present the integrated M-PCM–OFFD algorithm, and prove the main results on its performance in terms of (1) output mean estimation, (2) robustness to numerical truncation errors of simulators, and (3) estimation of the effects of the parameters. Section 4 includes simulation studies on some illustrative examples, and compares the performance of M-PCM–OFFD with existing function approximation approaches. Section 5 concludes the paper. For better readability, we move the proofs of all lemmas, theorems, and corollaries to the Appendix.

2. Preliminaries

We review fundamentals of the M-PCM and OFFD methods, which pave the foundation for main results in this paper.

2.1. M-PCM: Multivariate probabilistic collocation method

The Probabilistic Collocation Method (PCM) was developed to quickly and accurately estimate the output mean for a system of one uncertain input parameter, which may or may not be correlated [31,61]. M-PCM extends PCM to address systems of multiple uncertain input parameters [70,71]. Although computationally effective in each uncertainty dimension, it does not address the exponential expansion of computational costs with linearly increasing uncertainty dimensions. For a system mapping $g(x_1, x_2, \ldots x_m)$ of m uncertain input parameters, with the degree k_i of each parameter x_i up to $2n_i - 1$, a total number of $2^m \prod_{i=1}^{m} n_i$ simulations are needed to uniquely determine the mapping.

The procedure of M-PCM, Algorithm 1, shown at the next page, selects n_i simulation points for each parameter based on statistical knowledge of uncertain input parameters in general forms such as joint probabilistic distribution functions (pdfs), historical datasets, and simple low-order moments such as the mean and the variance. A total number of $\prod_{i=1}^{m} n_i$ simulations evaluated at combinations of these simulation points constructs a reduced-order mapping $g^+(x_1, x_2, \ldots x_m)$ with the degree of each parameter reduced to $n_i - 1$. When $m = 1$, we refer the method as PCM [31,61]. The reduced-order mapping correctly estimates the output mean of the original mapping as shown in Theorem 1. The proof can be found in [71]. In this paper, we assume the independence of these uncertain input parameters. Properties of the correlated case can be found in [71].

Theorem 1 (*[70,71]*)**.** *Consider a multivariate system with the following mapping*

$$g(x_1, x_2, \ldots x_m) = \sum_{k_1=0}^{2n_1-1} \sum_{k_2=0}^{2n_2-1} \cdots \sum_{k_m=0}^{2n_m-1} \Psi_{k_1,k_2,\ldots,k_m} \prod_{i=1}^{m} x_i^{k_i}, \tag{1}$$

where the coefficients $\Psi_{k_1,k_2,\ldots,k_m} \in R$, *and* $n_1, n_2, \ldots, n_m$ *are integers greater than* 1. *Assume that the uncertain parameters* $x_1, x_2, \ldots, x_m$ *follow independent distributions* $f_{X_1}(x_1)$, $f_{X_2}(x_2)$, $\ldots$, *and* $f_{X_m}(x_m)$ *respectively.* $g(x_1, x_2, \ldots x_m)$ *can be approximated by the following reduced-order system mapping*

$$g^+(x_1, x_2, \ldots x_m) = \sum_{k_1=0}^{n_1-1} \sum_{k_2=0}^{n_2-1} \cdots \sum_{k_m=0}^{n_m-1} \Omega_{k_1,k_2,\ldots,k_m} \prod_{i=1}^{m} x_i^{k_i}, \tag{2}$$

using the procedure shown in Algorithm 1, *where the coefficients* $\Omega_{k_1,k_2,\ldots,k_m} \in R$. *Then, the following equality holds*

$$E[g(x_1, x_2, \ldots x_m)] = E[g^+(x_1, x_2, \ldots x_m)].$$

Algorithm 1: M-PCM

/* Step 1: Choose M-PCM simulation set */

1 **for** $i \leftarrow 1$ **to** m **do**

2 Initialize $H_i^{-1}(x_i) = h_i^{-1}(x_i) = 0$ and $H_i^0(x_i) = h_i^0(x_i) = 1$;

3 **for** $k_i \leftarrow 1$ **to** n_i **do**

4 $H_i^{k_i}(x_i) = x_i h_i^{k_i-1}(x_i) - \langle x_i h_i^{k_i-1}(x_i), h_i^{k_i-1}(x_i) \rangle h_i^{k_i-1}(x_i) - \langle H_i^{k_i-1}(x_i), H_i^{k_i-1}(x_i) \rangle^{\frac{1}{2}} h_i^{k_i-2}(x_i)$;

5 $h_i^{k_i}(x_i) = H_i^{k_i}(x_i) / \langle H_i^{k_i}(x_i), H_i^{k_i}(x_i) \rangle^{\frac{1}{2}}$;

6 **end**

7 **end**

8 Find the roots of $h_i^{n_i}(x_i) = 0$ in Line 5 as the n_i PCM simulation, denoted as $x_{i(1)}, \cdots x_{i(n_i)}$;

/* Step 2: Run simulations */

9 For each m-tuple simulation point $(x_{1(r_1)}, x_{2(r_2)}, \ldots x_{m(r_m)})$ in the M-PCM simulation set, where $r_i \in \{1, \ldots n_i\}$, run simulation and find the corresponding output $g(x_{1(r_1)}, x_{2(r_2)}, \ldots x_{m(r_m)})$;

/* Step 3: Calculate the output mean */

10 Apply $\begin{bmatrix} a_{0,\ldots,0} \\ a_{0,\ldots,1} \\ \vdots \end{bmatrix} = \Gamma^{-1} \begin{bmatrix} g(x_{1(1)}, \ldots, x_{m(1)}) \\ g(x_{1(1)}, \ldots, x_{m(2)}) \\ \vdots \end{bmatrix}$, where $\Gamma =$

$$\begin{bmatrix} h_1^0(x_{1(1)})\ldots h_m^0(x_{m(1)}) & h_1^0(x_{1(1)})\ldots h_m^1(x_{m(1)}) & \cdots & h_1^{n_1-1}(x_{1(1)})\ldots h_m^{n_m-1}(x_{m(1)}) \\ h_1^0(x_{1(1)})\ldots h_m^0(x_{m(2)}) & h_1^0(x_{1(1)})\ldots h_m^1(x_{m(2)}) & \cdots & h_1^{n_1-1}(x_{1(1)})\ldots h_m^{n_m-1}(x_{m(2)}) \\ \vdots & \vdots & \ddots & \vdots \\ h_1^0(x_{1(n_1)})\ldots h_m^0(x_{m(n_m)}) & h_1^0(x_{1(n_1)})\ldots h_m^1(x_{m(n_m)}) & \cdots & h_1^{n_1-1}(x_{1(n_1)})\ldots h_m^{n_m-1}(x_{m(n_m)}) \end{bmatrix}$$

to find the coefficients $a_{k_1,k_2,\ldots k_m}$ in the reduced-order mapping: $g^+(x_1, x_2, \cdots x_m) = \sum_{k_1=0}^{n_1-1} \sum_{k_2=0}^{n_2-1} \cdots \sum_{k_m=0}^{n_m-1} a_{k_1,k_2,\ldots k_m} \prod_{i=1}^{m} h_i^{k_i}(x_i)$;

11 The predicted output mean is $a_{0,0,\ldots 0}$;

Notes of Algorithm 1

1. In Line 4, $\langle p(x_i), q(x_i) \rangle$ denotes the integration operation $\int p(x_i) q(x_i) f_{X_i}(x_i) dx_i$. $H_i^{k_i}(x_i)$ is the orthogonal polynomial of degree k_i for the uncertain input parameter x_i, and $h_i^{k_i}(x_i)$ is the normalized orthonormal polynomial [2].
2. In Line 8, the indices in the two column vectors are arranged in the lexicographic order, from $0, 0, \ldots 0$ to $n_1 - 1, n_2 - 1, \ldots n_m - 1$, and from $1, 1, \ldots 1$ to $n_1, n_2, \ldots n_m$, respectively. The entries in Γ are arranged accordingly.

Remarks

1. The coefficients $\Omega_{k_1,k_2,\ldots,k_m}$ in (2) can be derived by reorganizing terms in $g^+(x_1, x_2, \ldots x_m)$ obtained in Line 10 of Algorithm 1. The reduced-order mapping, as a by-product of this estimation process, enables further studies, such as sensitivity analysis and optimization [72,73].
2. Besides accurately predicting the output mean with a reduced computational cost, M-PCM has other attractive statistical characteristics [70,71]. In particular, it also precisely predicts the cross-statistics (i.e., statistics of cross input–output relationship) up to a certain degree [70,71]. Moreover, its performance is tightly connected to the minimum mean square estimation. These properties suggest that the reduced-order polynomial mapping $g^+(x_1, x_2, \ldots x_m)$ approximates the original system mapping $g(x_1, x_2, \ldots x_m)$ well over likely ranges of parameter values. We also note that the methodology can be modified to estimate high-order moments as well, but a larger τ may be required to achieve the same level of accuracy, due to the increase of the order of interactions. For instance, the second-order moment of the output can be estimated by evaluating the expected mapping of g^2, e.g., $E(g^2)$. We leave the study on high-order output statistics for future work.

3. In line 8 of Algorithm 1, Γ^{-1} can be calculated by the expression $(W\Gamma)^T W$, where $W = \text{diag}\{w_1, w_2, \ldots, w_l\}$, $w_i = 1/\|\Gamma_{i,:}\|$, and $l = \prod_{i=1}^{m} n_i$. $(\cdot)_{i,:}$ represents the i th row of the matrix in the parenthesis, and $\|\cdot\|$ is the L_2 norm. Computing polynomial coefficients from interpolation at the zeros of orthogonal polynomials is equivalent to computing these coefficients via Gauss quadrature (quadrature collocated at zeros of orthogonal polynomials). As $W\Gamma$ is an orthogonal matrix, and the linear system can be solved via simple transposition and matrix multiplication $(W\Gamma)^T W$ [6,17].
4. Instead of following lines 8–10 in Algorithm 1 to find mapping coefficients $\Omega_{k_1,k_2,\ldots,k_m}$, we can also directly calculate them using all M-PCM simulation points. Specifically, we can respectively replace each element $h_i^{k_i}(x_{i(r_i)})$ in matrix Γ with $x_i^{k_i}(x_{i(r_i)})$, to construct matrix L, where $x_i^{k_i}(x_{i(r_i)})$ represents the k_ith power of x_i evaluated at the PCM simulation point $x_{i(r_i)}$. The coefficients $\Omega_{k_1,k_2,\ldots,k_m}$ can be directly computed by

$$\begin{bmatrix} \Omega_{0,0,\ldots,0} \\ \Omega_{0,0,\ldots,1} \\ \vdots \\ \Omega_{n_1-1,n_2-1,\ldots,n_m-1} \end{bmatrix} = L^{-1} \begin{bmatrix} g(x_{1(1)}, x_{2(1)}, \ldots x_{m(1)}) \\ g(x_{1(1)}, x_{2(1)}, \ldots x_{m(2)}) \\ \vdots \\ g(x_{1(n_1)}, x_{2(n_2)}, \ldots x_{m(n_m)}) \end{bmatrix} \tag{3}$$

where

$$L = \begin{bmatrix} x_1^0(x_{1(1)})\ldots x_m^0(x_{m(1)}) & \cdots & x_1^{n_1-1}(x_{1(1)})\ldots x_m^{n_m-1}(x_{m(1)}) \\ x_1^0(x_{1(1)})\ldots x_m^0(x_{m(2)}) & \cdots & x_1^{n_1-1}(x_{1(1)})\ldots x_m^{n_m-1}(x_{m(2)}) \\ \vdots & \ddots & \vdots \\ x_1^0(x_{1(n_1)})\ldots x_m^0(x_{m(n_m)}) & \cdots & x_1^{n_1-1}(x_{1(n_1)})\ldots x_m^{n_m-1}(x_{m(n_m)}) \end{bmatrix}.$$

The indices in the two column vectors in (3) are arranged in the lexicographic order (see Note 2 of Algorithm 1 at the previous page). The output mean can then be calculated either through an integration

$$E[g(x_1, \ldots x_m)] = \int \cdots \int g^+(x_1, \ldots x_m) f_{X_1}(x_1) \cdots f_{X_m}(x_m) dx_1 \cdots dx_m, \tag{4}$$

or through a simple transformation

$$E[g(x_1, \ldots x_m)] = (\Gamma^{-1} L)_{1,:} B, \tag{5}$$

where B is the coefficient vector on the left-hand side of (3). The orthogonal basis (captured by matrix Γ) is preferred over the natural basis (matrix L) to achieve better invertability, as captured by the condition number metric [71]. In this paper, we use L for the ease of presentation in proofs; however the results also apply to the case when Γ is used.

Even though it provides a significant computational load reduction from $2^m \prod_{i=1}^{m} n_i$ to $\prod_{i=1}^{m} n_i$, M-PCM does not scale with the dimension m of uncertain input parameters. In realistic applications, high-order cross-terms in a mapping have negligible effects on the output [4]. With such assumptions that generally hold for real-world large-scale complex system applications, only a subset of the M-PCM simulation set is needed to perform computation-intensive simulations. This motivates our study that leverages OFFD to address the curse of dimensionality problem of M-PCM.

2.2. OFFDs: Orthogonal fractional factorial designs

OFFDs have been widely used to reduce the number of experiments and to enact various parameter and system studies such as influential parameter identification, sensitivity analysis, model optimization and software testing [38,42,64]. For an experiment that involves multiple parameters (or factors) with each parameter evaluated at several values (or levels), all combinations of the parameter values form a large design space, called the *full factorial design*. OFFDs select a subset from the full factorial design to retain the significant effects of input parameters (i.e., the main effects of single parameters and low-order interaction effects of a few parameters) with minimal aliases. To illustrate the concepts of main effect, interaction effect, and aliases, we here describe the 2-level full factorial design and OFFD designs (for which each parameter only takes two values [4,22,48,49]) from the estimation perspective.

	x_{c_1}	x_{c_2}	x_{c_3}	$x_{c_1}x_{c_2}$	$x_{c_1}x_{c_3}$	$x_{c_2}x_{c_3}$	$x_{c_1}x_{c_2}x_{c_3}$	y
1	−1	−1	−1	+1	+1	+1	−1	y_1
2	+1	−1	−1	−1	−1	+1	+1	y_2
3	−1	+1	−1	−1	+1	−1	+1	y_3
4	+1	+1	−1	+1	−1	−1	−1	y_4
5	−1	−1	+1	+1	−1	−1	+1	y_5
6	+1	−1	+1	−1	+1	−1	−1	y_6
7	−1	+1	+1	−1	−1	+1	−1	y_7
8	+1	+1	+1	+1	+1	+1	+1	y_8

(a)

	x_{c_1}	x_{c_2}	x_{c_3}	y
1	−1	−1	+1	y_5
2	+1	−1	−1	y_2
3	−1	+1	−1	y_3
4	+1	+1	+1	y_8

	x_{c_1}	x_{c_2}	x_{c_3}	y
1	−1	−1	−1	y_1
2	+1	+1	−1	y_4
3	−1	+1	+1	y_7
4	+1	−1	+1	y_6

(b)

Fig. 1. Illustration of 3-factor (a) 2^3 full factorial design and (b) the design tables of two 2^{3-1}_{III} OFFDs.

Consider m parameters, $x_1, x_2, \ldots, x_m$, coded as factors, $x_{c_1}, x_{c_2}, \ldots, x_{c_m}$, each of which is evaluated at two levels (or coded values), '+1' and '−1'. The 2-level full factorial design of 2^m parameter combinations constructs a saturated model that describes the *effects* of these m factors on the output [49]:

$$y = \beta_0 + \sum_{i=1}^{m} \beta_i x_{c_i} + \sum_{i=1}^{m-1} \sum_{j=i+1}^{m} \beta_{ij} x_{c_i} x_{c_j} + \sum_{i=1}^{m-2} \sum_{j=i+1}^{m-1} \sum_{k=j+1}^{m} \beta_{ijk} x_{c_i} x_{c_j} x_{c_k} + \cdots + \beta_{1\ldots m} x_{c_1} x_{c_2} \ldots x_{c_m} + \epsilon \tag{6}$$

where the coefficients reflect the effects of corresponding terms to the output, and ϵ is a random noise with zero mean. Denoting '$\bar{}$' as the average operator, the standard least squares estimation gives the estimates $\hat{\beta}_0 = \bar{y}$, $\hat{\beta}_i = \frac{1}{2}(\bar{y}_{i=+} - \bar{y}_{i=-})$, $\hat{\beta}_{ij} = \frac{1}{2}(\bar{y}_{ij=+} - \bar{y}_{ij=-})$, and $\hat{\beta}_{ij\ldots} = \frac{1}{2}(\bar{y}_{ij\ldots=+} - \bar{y}_{ij\ldots=-})$. Here, $\bar{y}$ is the mean of all outputs, $\bar{y}_{i=\pm}$ denote the output means when $x_{c_i} = \pm 1$, and $\bar{y}_{ij\ldots=\pm}$ denote the output means when $x_{c_i} x_{c_j} \ldots = \pm 1$. Defining the *main effect* χ_i of factor x_{c_i} as $\chi_i = \bar{y}_{i=+} - \bar{y}_{i=-}$, and the *interaction effect* (interaction of multiple factors) $\chi_{ij\ldots}$ of cross-term $x_{c_i} x_{c_j} \ldots$ as $\chi_{ij\ldots} = \bar{y}_{ij\ldots=+} - \bar{y}_{ij\ldots=-}$, we obtain the least squares estimates of $\hat{\beta}_i = \frac{1}{2}\chi_i$, and $\hat{\beta}_{ij\ldots} = \frac{1}{2}\chi_{ij\ldots}$ [4,49]. Full factorial designs estimate all main and interaction effects (of any order) independently of one another. The disadvantage is that they are expensive in requiring many simulation runs. A full factorial design for three factors is shown in Fig. 1(a).

A 2-level OFFD is described as $2^{m-\gamma}_R$. The *fractionation constant*, γ, in the range of $1 \le \gamma \le m - \lceil \log_2(m+1) \rceil$ [4], indicates that $N = 2^{m-\gamma}$ simulations are selected from the full set of 2^m simulations [22,48]. The upper bound of γ is determined by the minimum number of runs ($N = m + 1$) to estimate m main effects and the mean. The *resolution* R, represented by a Roman numerical subscript, describes the length of the shortest *generator*, which defines the rules to generate factor levels [48]. For example, the two halved-size OFFDs in Fig. 1(b) are both 2^{3-1}_{III} designs produced by setting $x_{c_3} = \pm x_{c_1} x_{c_2}$. The corresponding generator is $I = \pm x_{c_1} x_{c_2} x_{c_3}$, and the resolution $R = III$.

A 2-level OFFD is captured by the *design table*, where each entry i, j represents the level of factor j selected in the i th simulation. Now let us briefly illustrate the procedure to generate $2^{m-\gamma}_R$ OFFDs for given m and γ. The steps in Algorithm 2 illustrated below deviate to some extent from the iterative procedures in the standard experiment design literature [16] to address the uncertainty evaluation problem.

Algorithm 2: OFFD

/* Step 1: Generate the $2^{m-\gamma}$ full factorial design for $m - \gamma$ factors. */

1 List all $2^{m-\gamma}$ combinations for the $m - \gamma$ factors;

/* Step 2: Specify γ generators. */

2 Select generators that maximize the resolution R;

/* Step 3: Determine the levels of all other γ factors. */

3 Generate the levels for all other γ factors using the generators selected in Step 2;

In Line 2, we refer to [60] for standard generator designs.

Lemma 1. *A $2^{m-\gamma}_R$ OFFD satisfies $m - \gamma \ge R - 1$.*

The $2_R^{m-\gamma}$ OFFDs are characterized by the *balance* and *orthogonality* properties [48]. Denote the OFFD design table as D of dimension $2^{m-\gamma} \times m$, and its i th column as $\mathbf{d}_i$. Then the balance property leads to $\mathbf{d}_i \cdot \mathbf{1} = 0$, where $1 \le i \le m$, and '$\cdot$' denotes the inner product of two vectors. The orthogonality property leads to $\mathbf{d}_i \cdot \mathbf{d}_j = 0$, where $1 \le i, j \le m$ and $i \ne j$.

An advantage of OFFDs is that they are inexpensive, but with the "price-paid" that every fractional design – orthogonal or otherwise – has *confounding (aliasing)* by which main effect estimates reflect the factor of interest, but also have contributions from other sources (namely, interaction effect(s)) [4,49]. The orthogonality and balance properties of OFFDs permit the confounding to be isolated, pre-determined (pre-data collection), and minimized in the sense that main effect estimates are confounded with high-order interaction effects and not with other main effects [48]. When these higher-order interaction effects are negligible, OFFDs provide very good estimates of main and interaction effects.

Multiple $2^{m-\gamma}$ OFFDs may exist with different confounding structures [22,48]. For given m and γ, it is desirable to choose the highest possible resolution R, because this assures that the main effects are only confounded with the highest-order interaction effects that can be achieved [4]. For designs of $\gamma > 1$, the main effect of factor x_i is only confounded with the interaction effects of at least $R - 1$ factors among the other $m - 1$ factors (with $R \le m$). Moreover, the k-factor interaction effect is only confounded with the interaction effects of at least $R - k$ factors (when $k \le R - k$) [48]. For instance, if 2-factor interaction effects cannot be neglected, a design of resolution $R = III$ may not be sufficient. At the extreme, consider the *half-fraction design* (i.e., $\gamma = 1$). The generator for 2_m^{m-1} has a form of $I = \pm x_{c_1} x_{c_2} \cdots x_{c_m}$. Under this resolution, the main effect of factor x_i is only confounded with the interaction effect of all other $m - 1$ factors, and a k-factor interaction effect is only confounded with the interaction effects of all other $m - k$ factors (assuming $k \le m - k$). Therefore, if $(m - k)$-factor interaction effects can be negligible, the k-factor interaction effect estimates are also very accurate.

In viewing the M-PCM simulation set as a full factorial design, Algorithm 2 provides a systematic procedure to select a subset of m-tuple simulation points to reduce the number of simulations. In the next section, we show that the resulting reduced-order polynomial mapping correctly estimates the output mean of the original system mapping, and is the most robust to numerical truncation errors of simulators.

3. Integrated M-PCM–OFFD design and analysis

In this section, we integrate M-PCM and OFFD to break the curse of dimensionality for effective output mean estimation. We first present the integrated M-PCM–OFFD algorithm, and then analyze its performance using three metrics: (1) estimation of output mean, (2) robustness to numerical truncation errors of simulators, and (3) estimation of significant effects. The specific OFFD to select is based on the knowledge that cross-terms in the original system mapping involve at most a certain number of parameters. Such knowledge can be obtained from experimental studies of underlying physical systems.

3.1. The M-PCM–OFFD algorithm

Consider an original system mapping of m uncertain input parameters, $x_1, x_2, \ldots, x_m$, each with a degree up to 3:

$$g(x_1, x_2, \ldots x_m) = \sum_{k_1=0}^{3} \sum_{k_2=0}^{3} \cdots \sum_{k_m=0}^{3} \psi_{k_1,k_2,\ldots,k_m} \prod_{i=1}^{m} x_i^{k_i}, \tag{7}$$

where the coefficients $\psi_{k_1,\ldots,k_m} \in R$. Assume that the uncertain input parameters follow independent distributions $f_{X_1}(x_1)$, $f_{X_2}(x_2)$, $\ldots$, and $f_{X_m}(x_m)$. In addition, assume that cross-terms have up to τ parameters, where τ, the *maximal cross-product power*, is an integer of $1 \le \tau \le m$. In other words, $\psi_{k_1,\ldots,k_m} = 0$ if more than τ of $k_1, \ldots k_m$ are non-zero. τ is determined either based on prior field knowledge of the physical system, or through an iterative procedure that trades off between the computational cost and the estimation accuracy.

Algorithm 3 constructs a reduced-order polynomial mapping

$$g^*(x_1, x_2, \ldots x_m) = \sum_{k_1=0}^{1} \sum_{k_2=0}^{1} \cdots \sum_{k_m=0}^{1} \Omega_{k_1,k_2,\ldots,k_m} \prod_{i=1}^{m} x_i^{k_i}, \tag{8}$$

where the coefficients $\Omega_{k_1,k_2,\ldots,k_m} \in R$, and $\Omega_{k_1,k_2,\ldots,k_m} = 0$ if more than τ of $k_1, k_2, \ldots k_m$ are non-zero. Note that $\tau = \max \sum_i k_i$ in the reduced-order mapping g^*, where $i \in \{1, 2, \ldots m\}$.

Algorithm 3: M-PCM–OFFD

/* Step 1: Select the M-PCM simulation set */

1 Follow Step 1 of the Algorithm 1 where $n_i = 2$ for all $i \in \{1, 2, \ldots m\}$;

2 **if** $m > 2$ *and* $1 \le \tau \le \lceil \frac{m}{2} \rceil - 1$ **then**

3 Go to Step 2;

4 **else**

5 Go to Step 4, as no reduction of simulations can be achieved;

6 $l_{\text{offd}} \leftarrow 2^m$, where l_{offd} is the number of M-PCM–OFFD simulations;

7 **end**

/* Step 2: Determine γ^* to save $2^m - 2^{m-\gamma^*}$ simulations. */

8 $\gamma^* \leftarrow \max\{\gamma \mid 1 \le \gamma \le m - \lceil \log_2 (\sum_{i=0}^{\tau} \binom{m}{i}) \rceil$, *and* $2^{m-\gamma}_R$ *OFFD exists with* $R \ge 2\tau + 1\}$;

/* Step 3: Select the M-PCM-OFFD simulation set using the $2^{m-\gamma^*}_R$ OFFD. */

9 Use Algorithm 2 to select $l_{\text{offd}} = 2^{m-\gamma^*}$ parameter combinations to form m-tuple simulation points from the full set of 2^m M-PCM simulation points obtained in Step 1. These $2^{m-\gamma^*}$ m-tuple simulation points constitute the *M-PCM–OFFD simulation set*;

/* Step 4: Simulation */

10 Run simulation for the M-PCM–OFFD simulation set at each of the $2^{m-\gamma^*}_R$ m-tuple simulation points;

/* Step 5: Calculate the output mean */

11 **if** $l = l_{offd}$ **then**

12 Find the coefficients in (8) similar to Step 3 of Algorithm 1, but with a reduced-size matrix Γ, denoted as the *input matrix* $\Gamma'_{\text{offd}} \in R^{l_{\text{offd}} \times l}$, which excludes those columns with more than τ of $k_1, k_2, \ldots, k_m$ being nonzero, and rows representing points not selected in the reduced M-PCM–OFFD simulation set;

13 **else**

14 Find the coefficients in (8) by replacing $\Gamma'^{-1}_{\text{offd}}$ with $(\Gamma'^{T}_{\text{offd}} \Gamma'_{\text{offd}})^{-1} \Gamma'^{T}_{\text{offd}}$ [43].

15 **end**

16 The predicted output mean is $a_{\underbrace{0,\ldots,0}_{m}}$.

Notes of Algorithm 3

1. In Line 8, $l = \sum_{i=0}^{\tau} \binom{m}{i}$ denotes the number of M-PCM–OFFD coefficients to be estimated.
2. Step 4 is the most time-consuming step for large-scale complex system applications.

Remarks

1. In Line 16 of Algorithm 3, the output mean can also be calculated by integrating the reduced-order mapping using (4), or through a simple matrix manipulation. In particular, $E[g(x_1, x_2, \ldots x_m)] = [(\Gamma'^{-1}_{\text{offd}} L'_{\text{offd}})_{1,:}] B'$ if $l_{\text{offd}} = l$, or $E[g(x_1, x_2, \ldots x_m)] = [((\Gamma'^{T}_{\text{offd}} \Gamma'_{\text{offd}})^{-1} L'_{\text{offd}})_{1,:}] B'$ if $l_{\text{offd}} > l$, where $L'_{\text{offd}} \in R^{l_{\text{offd}} \times l}$ is a reduced-size L matrix derived in the way same as matrix Γ'_{offd}. B' is a reduced-size B vector, which excludes entries corresponding to cross-terms of more than τ parameters.
2. The ordering of entries in matrix L'_{offd} or Γ'_{offd} does not need to strictly follow Note 2 of Algorithm 1. They only need to match the orderings of simulation points and the simulated outputs. Without loss of generality (WLOG), we assume in the sequel that columns in matrix L'_{offd} are arranged in the graded reverse lexicographic order as follows for the ease of proofs. First, the number of parameters in these columns increases from 0 to τ. Second, columns of the same number of parameter are arranged in the order of increasing indices. For instance, for a system with $m = 3$ parameters and the maximal cross-product power $\tau = 3$, the columns in L'_{offd} are arranged in the following order: $\{1, x_1, x_2, x_3, x_1x_2, x_1x_3, x_2x_3, x_1x_2x_3\}$.

In this sequel, we only analyze the case of using matrix L'_{offd} to construct $g^*(x_1, x_2, \ldots x_m)$ for the ease of presentation. As matrices Γ'_{offd} and L'_{offd} share the same structure, similar conclusions hold if using Γ'_{offd} instead.

3.2. Performance analysis on estimation correctness

In this section, we consider the case where a further reduction of simulations is possible (i.e., when $1 \le \tau \le \lceil \frac{m}{2} \rceil - 1$, and $m > 2$), and show that the reduced M-PCM–OFFD simulation set obtained using Algorithm 3 correctly estimates the output mean of the original system mapping with the degree of each parameter up to 3 (Theorem 2). We first present several lemmas. We show in Lemma 2 that the reduced-order mapping does not increase the maximal cross-product power.

Lemma 2. *Consider an original system mapping $g(x_1, x_2, \ldots x_m)$ of maximal cross-product power τ, (7). The maximal cross-product power of the reduced-order mapping $g^+(x_1, x_2, \ldots x_m)$ is also τ.*

Given the maximal cross-product power τ in cross-terms of $g(x_1, x_2, \ldots x_m)$, or equivalently $g^+(x_1, x_2, \ldots x_m)$ according to Lemma 2, Lemma 3 partially justifies Steps 1 and 2 of Algorithm 3.

Lemma 3. *Consider the reduced-order M-PCM mapping $g^+(x_1, x_2, \ldots x_m)$, (2), of maximal cross-product power τ. An OFFD design can further reduce the number of simulations, if $1 \le \tau \le \lceil \frac{m}{2} \rceil - 1$, and $m > 2$ (Step 1 of Algorithm 3). The number of simulations can be reduced from 2^m to $2^{m-\gamma_{\max}}$, using a $2^{m-\gamma_{\max}}$ OFFD, where $\gamma_{\max} = m - \lfloor \log_2 \left(\sum_{i=0}^{\tau} \binom{m}{i} \right) \rfloor$.*

In order for the $2^{m-\gamma_{\max}}$ OFFD in Lemma 3 to lead to a valid M-PCM–OFFD design (e.g., Step 5 of Algorithm 3 is possible), the input matrix L'_{offd} needs to have full column rank. Lemmas 4–6 state that this requirement is satisfied when the resolution $R \ge 2\tau + 1$. We first introduce a matrix $V \in R^{l_{\text{offd}} \times l}$, and establish its connection with the $2^{m-\gamma^*}_R$ design table, D, in Lemma 4. Here we denote '$\circ$' as the hadamard product.

Lemma 4. *Consider a matrix $V \in R^{l_{offd} \times l}$ constructed from the input matrix $L'_{offd} \in R^{l_{offd} \times l}$ (with $1 \le \tau \le \lceil \frac{m}{2} \rceil - 1$ and $m > 2$). In matrix V, each column $\mathbf{v}_i$ is the i th column of L'_{offd} with its entries $x_{k(1)}$ and $x_{k(2)}$ replaced by the coded factors '-1' and '$+1$' respectively, where $k \in S_i$ is the index of input parameters, and S_i is a set that includes all the indices of input parameters in column i. The following equalities hold. $\mathbf{v}_1 = \mathbf{1}$. The column $\mathbf{v}_i$ in matrix V, where $1 < i \le m + 1$, equals the $(i-1)$th column of the OFFD design table D, i.e., $\mathbf{v}_i = \mathbf{d}_{i-1}$. Column $\mathbf{v}_i$, where $i > m + 1$, equals $\prod_{k \in S_i} \mathbf{v}_k$.*

The connection of V and the OFFD design table shown in Lemma 4 suggests that V can also be directly constructed from the design table. Lemmas 1, 2 and 4 lead to Lemma 5, which states the condition for the orthogonality of matrix V.

Lemma 5. *Any two columns of V are orthogonal, i.e., $\mathbf{v}_i \cdot \mathbf{v}_j = 0$, for any $i \neq j$, and $i, j = \{1, 2, \ldots l\}$, if $R \ge 2\tau + 1$.*

Based on Lemma 5, Lemma 6 shows that the matrix L'_{offd} constructed by M-PCM–OFFD is of full column rank (i.e., $\text{rank}(L'_{\text{offd}}) = l$).

Lemma 6. *When $1 \le \tau \le \lceil \frac{m}{2} \rceil - 1$ and $m > 2$, the input matrix $L'_{offd} \in R^{l_{offd} \times l}$ (constructed using Algorithm 3) is of full column rank, and can be expressed as $L'_{offd} = VU$, where $U \in R^{l \times l}$ is an upper triangular matrix, with its i th diagonal entry*

$$U_{i,i} = \begin{cases} 1 & \text{if } i = 1 \\ \frac{1}{2^{\xi_i}} \prod_{k \in S_i} \Delta x_k & \text{if } i \neq 1 \end{cases} \tag{9}$$

where $\Delta x_k = x_{k(2)} - x_{k(1)}$, ξ_i is the size of S_i, and $x_{k(2)} > x_{k(1)}$ WLOG.

Theorem 1 and Lemmas 2–6 lead to the following theorem, which states that M-PCM–OFFD correctly estimates the output mean of the original system mapping.

Theorem 2. *When* $1 \le \tau \le \lceil \frac{m}{2} \rceil - 1$ *and* $m > 2$, *the reduced-order mapping* $g^*(x_1, x_2, \ldots x_m)$, (8), *constructed using the integrated M-PCM and* $2^{m-\gamma^*}$ *OFFD methods (Algorithm 3) correctly estimates the output mean of the original system mapping* $g(x_1, x_2, \ldots x_m)$, *i.e.,*

$$E[g(x_1, x_2, \ldots x_m)] = E[g^*(x_1, x_2, \ldots x_m)]. \tag{10}$$

The algorithm reduces the number of simulations from 2^m *to* $2^{m-\gamma^*}$ *in the range of* $\left[2^{\lceil \log_2(m+1) \rceil}, 2^{m-1}\right]$.

3.3. Performance analysis on the robustness to numerical truncation errors of simulators

In this section, we study the robustness of M-PCM–OFFD to numerical truncation errors of simulators. A robustness metric and problem formulation are described in Section 3.3.1. The optimality of M-PCM–OFFD under this metric is proved in Section 3.3.2.

3.3.1. Metric and problem formulation

The M-PCM–OFFD algorithm involves L'^{-1}_{offd} or $\left(L'^{T}_{\text{offd}} L'_{\text{offd}}\right)^{-1} L'^{T}_{\text{offd}}$, which requires L'_{offd} to have full column rank. We showed in Lemma 6 that this full column property holds when the resolution satisfies $R \ge 2\tau + 1$. Numerical truncation errors caused by allowable parameter resolutions of simulation software [71] and computational limitations of computing devices [33] may introduce disturbances and push L'_{offd} to lose rank. Even if such an L'_{offd} under disturbances does not lose rank, the correctness of inversion computation becomes sensitive to such numerical truncation errors [13]. As such, L'_{offd} needs to have a *large margin to rank loss* under parameter value disturbances for practical use.

We here show that L'_{offd} has the largest margin to rank loss among all subsets of the same size selected from the M-PCM simulation set. To facilitate the analysis, we introduce L' to represent a matrix constructed in the similar way as L'_{offd}, by using an arbitrary subset of the size $2^{m-\gamma^*}$ in the M-PCM simulation set. We also define a metric to measure the distance to column rank loss as the *full-column-rank margin* $M(L')$, which is measured by the Frobenius norm [35] of the smallest perturbation matrix to make L' lose rank [12,33],

$$M(L') = \min\left\{ \|e\|_F \mid \text{rank}(L' + e) < l \right\} \tag{11}$$

where $e \in R^{l_{\text{offd}} \times l}$ is a perturbation matrix on parameter values.

3.3.2. Optimality analysis

Lemma 7 shows the expression of full-column-rank margin for M-PCM–OFFD when the maximal cross-product power $\tau = 1$. Theorem 3 shows that OFFD produces the largest full-column-rank margin among all simulation subsets of the same size selected from the M-PCM simulation set. Lemma 8 and Theorem 4 extend the results to the general case where $1 \le \tau \le \lceil \frac{m}{2} \rceil - 1$. Two corollaries follow to show exemplary designs where L' loses full column rank and therefore cannot be used for output mean estimation.

Lemma 7. *Consider an original system mapping* $g(x_1, x_2, \ldots x_m)$, (7), *with* $m > 2$ *and* $\tau = 1$. *The M-PCM–OFFD simulation set (selected using Algorithm 3) produces the full-column-rank margin:*

$$M(L'_{offd}) = \frac{\sqrt{l_{offd}}}{2} \min\{\Delta x_1, \Delta x_2, \ldots \Delta x_m\} \tag{12}$$

Theorem 3. *Consider an original system mapping* $g(x_1, x_2, \ldots x_m)$, (7), *with* $m > 2$ *and* $\tau = 1$. *The M-PCM–OFFD simulation set (selected using Algorithm 3) has the largest full-column-rank margin among all subsets of* $2^{m-\gamma^*}$ *simulation points in the M-PCM simulation set of size* 2^m. *Mathematically,* $\max\left(M(L')\right) = M(L'_{offd})$.

Before we show Lemma 8 and Theorem 4, let us first introduce some notations. Define $A = \{1, 2, \ldots m\}$, and $A_k = A - \{k\}$, where $k \in A$. We construct a set $A_{k,i}$ from A_k to contain all sets consisting i number of elements in A_k. The size of $A_{k,i}$ is $\binom{m-1}{i}$. The j th set in $A_{k,i}$ is denoted as $A_{k,i,j}$. For instance, when $k = 1, i = 2, j = 2$ and $m = 4$, $A = \{1, 2, 3, 4\}$, $A_1 = \{2, 3, 4\}$, $A_{1,2} = \{\{2, 3\}, \{2, 4\}, \{3, 4\}\}$, and $A_{1,2,2} = \{2, 4\}$.

Lemma 8. *Consider an original system mapping* $g(x_1, x_2, \ldots x_m)$, (7), *with* $m > 2$ *and* $1 \le \tau \le \lceil \frac{m}{2} \rceil - 1$. *The M-PCM-OFFD simulation set (selected using Algorithm 3) produces the full-column-rank margin:*

$$M(L'_{offd}) = \frac{\sqrt{l_{offd}}}{2} \min_{k \in A} \left\{ \Delta x_k \sqrt{1 + \sum_{i=1}^{\tau} \left\lfloor \frac{1}{2^i} \sum_{j=1}^{m-1} \prod_{a \in A_{k,i,j}} \left(x_{a(1)}^2 + x_{a(2)}^2 \right) \right\rfloor} \right\}, \quad (13)$$

Theorem 4. *Consider an original system mapping* $g(x_1, x_2, \ldots x_m)$, (7), *with* $m > 2$ *and* $1 \le \tau \le \lceil \frac{m}{2} \rceil - 1$. *The M-PCM-OFFD simulation set (selected using Algorithm 3) has the largest full-column-rank margin among all subsets of* $2^{m-\gamma^*}$ *simulation points in the M-PCM simulation set of size* 2^m. *Mathematically,* $\max(M(L')) = M(L'_{offd})$.

Corollaries 1 and 2 list two sufficient conditions for L' to be singular, i.e., $M(L') = 0$.

Corollary 1. *Consider an original system mapping* $g(x_1, x_2, \ldots x_m)$, (7), *with* $m > 2$ *and the maximal cross-product power* $\tau = 1$. *A design of* $2^{m-\gamma^*}$ *simulation points selected from the M-PCM simulation set of size* 2^m *leads to* $M(L') = 0$, *if there exist two columns in* L' *that contain a pair of input parameters with no more than two out of four combinations of levels.*

Corollary 2. *Consider an original system mapping* $g(x_1, x_2, \ldots x_m)$, (7), *with* $m > 2$ *and* $1 < \tau \le \lceil \frac{m}{2} \rceil - 1$. *A design of* $2^{m-\gamma^*}$ *simulation points selected from the M-PCM simulation set of size* 2^m *leads to* $M(L') = 0$, *if there exist two columns in* L' *which contain a pair of input parameters with no more than three out of four combinations of levels.*

3.4. Performance analysis of the correctness of effect estimation

For the completeness of performance analysis, we show in Theorem 5 the performance of M-PCM–OFFD in estimating effects of input parameters on the output.

Theorem 5. *Consider an original m-parameter system mapping* $g(x_1, x_2, \ldots x_m)$, (7), *The M-PCM–OFFD simulation set (selected using Algorithm 3) correctly estimates the main effects and all interaction effects.*

4. Illustrative examples and comparative studies

We first use a 3-parameter small example to illustrate the whole design procedures and properties of the integrated M-PCM–OFFD algorithm, and then briefly discuss a 50-parameter example to show its effectiveness in larger-size problems. To further demonstrate the performance of M-PCM–OFFD, we compare it with existing approaches reviewed in Section 1.

4.1. A small-scale example

Consider an original system mapping $g(x_1, x_2, x_3) = 1 + x_1 + x_1^2 + x_1^3 + x_2 + x_2^2 + x_2^3 + x_3 + x_3^2 + x_3^3$, where x_1 follows an exponential distribution $f_{X_1}(x_1) = 2e^{-2x_1}$, $x_1 \ge 0$, x_2 follows a uniform distribution of $f_{X_2}(x_2) = \frac{1}{15}$, $5 \le x_2 \le 20$, and x_3 also follows a uniform distribution of $f_{X_3}(x_3) = \frac{1}{5}$, $5 \le x_3 \le 10$. The output mean is $E[g(x_1, x_2, x_3)] = \int\int\int g(x_1, x_2, x_3) f_{X_1}(x_1) f_{X_2}(x_2) f_{X_3}(x_3) dx_1 dx_2 dx_3 = 3381.1$. Identifying all coefficients requires $4^3 = 64$ simulations.

Now let us use the M-PCM–OFFD algorithm (Algorithm 3) to choose only 4 simulations. First, we choose the M-PCM simulation set of 8 points based upon the pdf of each parameter: $p_1 = (x_{1(1)}, x_{2(1)}, x_{3(1)})$, $p_2 = (x_{1(2)}, x_{2(1)}, x_{3(1)})$, $p_3 = (x_{1(1)}, x_{2(2)}, x_{3(1)})$, $p_4 = (x_{1(2)}, x_{2(2)}, x_{3(1)})$, $p_5 = (x_{1(1)}, x_{2(1)}, x_{3(2)})$, $p_6 = (x_{1(2)}, x_{2(1)}, x_{3(2)})$, $p_7 = (x_{1(1)}, x_{2(2)}, x_{3(2)})$, and $p_8 = (x_{1(2)}, x_{2(2)}, x_{3(2)})$, where $x_{1(1)} = 0.2929$, $x_{1(2)} = 1.7071$, $x_{2(1)} = 8.1699$, $x_{2(2)} = 16.8301$, $x_{3(1)} = 6.0566$ and $x_{3(2)} = 8.9434$. We then use the 2_{III}^{3-1} OFFD (as the design table shows in Fig. 1(b)) to select the M-PCM–OFFD simulation set $\{p_2, p_3, p_5, p_8\}$ (as the 3-D cube shows in Fig. 2(a)) or $\{p_1, p_4, p_6, p_7\}$. The input

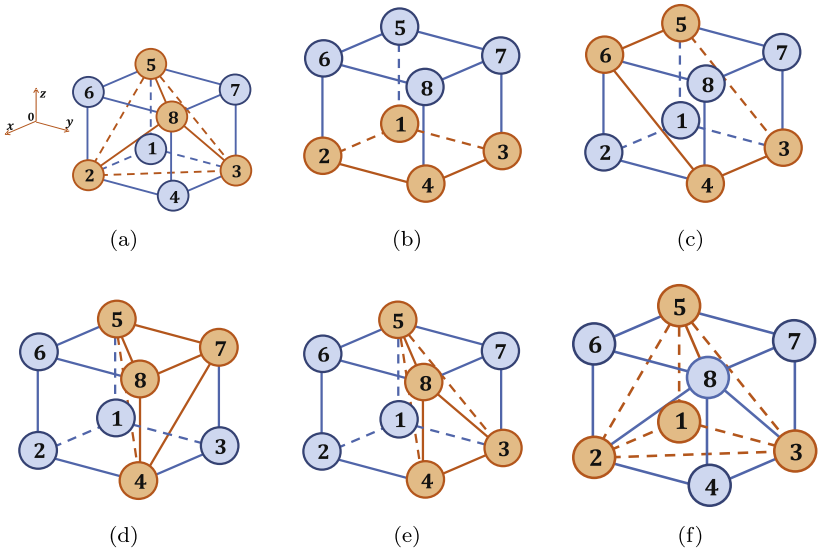

Fig. 2. 3-D cube representation of the (a) 2^{3-1}_{III} OFFD, and (b)-(f) other possible simulation sets of size 4 in the M-PCM simulation set of size 8. Each vertex i of the 3-D cube represents the M-PCM simulation point p_i. The selected subset is marked in orange.

matrix for the first design is $L'_{\text{offd}} = \begin{bmatrix} 1 & x_{1(1)} & x_{2(1)} & x_{3(2)} \\ 1 & x_{1(2)} & x_{2(1)} & x_{3(1)} \\ 1 & x_{1(1)} & x_{2(2)} & x_{3(1)} \\ 1 & x_{1(2)} & x_{2(2)} & x_{3(2)} \end{bmatrix}$. We then run simulations to evaluate $g(x_1, x_2, x_3)$ at these 4 M-PCM–OFFD simulation points and estimate coefficients of the reduced-order polynomial mapping $g^*(x_1, x_2, x_3) = -4442.2 + 6.5x_1 + 513.5x_2 + 186.8x_3$.

The output mean of $g^*(x_1, x_2, x_3)$ is $E[g^*(x_1, x_2, x_3)] = \int\int\int g^*(x_1, x_2, x_3) f_{X_1}(x_1) f_{X_2}(x_2) f_{X_3}(x_3) dx_1 dx_2 dx_3 = 3381.1$, same as the original output mean. Compared with M-PCM–OFFD that requires only 4 simulation points to generate the accurate output mean, M-PCM requires 8 simulation points, and the MC requires about 2400 simulations to reach the range from 3231.1 to 3531.1 (which is ± 150 around the true mean 3381.1) with 97% confidence.

To verify the robustness property of the integrated algorithm to numerical truncation errors of simulators, we calculate the full-column-rank margin of L'_{offd}, and compare it with other designs. According to Lemma 7, we find $M(L'_{\text{offd}}) = \min\{\Delta x_1, \Delta x_2, \Delta x_3\} = 1.4142$, where $\Delta x_1 = 1.4142$, $\Delta x_2 = 8.6602$ and $\Delta x_3 = 2.8868$. The minimal $\|e\|_F$ is achieved when the perturbation errors $e_{x_{1(1)}} = 0.7071$, $e_{x_{1(2)}} = -0.7071$, and $e_{x_{i(1)}} = e_{x_{i(2)}} = 0$ for all $i' = 1$.

Other possible subsets of simulation points include (1) four points on one surface (Fig. 2(b)), (2) four points on the diagonal plane (Fig. 2(c)), (3) three points on two surfaces (Fig. 2(d)), (4) three points on only one surface (Fig. 2(e)), and (5) three points on three surfaces (Fig. 2(f)). The full-column-rank margins of the input matrices for these designs take one of the four values $\{0, 0.8660, 1.2247, 1.4142\}$. The M-PCM–OFFD simulation set is thus the most robust to numerical truncation errors of simulators, with the largest full-column-rank margin 1.4142.

4.2. A large-scale example

In this example, we consider a system of 50 uncertain input parameters. The original system mapping $g(x_1, x_2, \ldots x_{50}) = \sum_{i=1}^{50} x_i + \sum_{i=1}^{50} x_i^2 + \sum_{i=1}^{50} x_i^3$, where each parameter x_i follows a uniform distribution of $f_{X_i}(x_i) = 1, 0 \le x_i \le 1$. The output mean is $E[g(x_1, x_2, \ldots x_{50})] = 55.1667$.

To obtain the accurate output mean, M-PCM requires 2^{50} simulations. Assuming that $\tau = 1$, we have $y^* = m - \lceil \log_2(1 + 50) \rceil = 44$ and $R = 3 \ge 2\tau + 1$, and we can further use the 2^{50-44}_{III} OFFD to reduce the size of simulation set from 2^{50} to 64. The reduced-order polynomial mapping estimated using M-PCM–OFFD is $g^*(x_1, x_2, \ldots x_{50}) = -15.6667 + 2.8333 \sum_{i=1}^{50} x_i$, which correctly estimates the output mean $E[g^*(x_1, x_2, \ldots x_{50})] = 55.1667$. This example further shows the effectiveness of M-PCM–OFFD in evaluating statistical system performance, especially for systems of high-dimensional uncertain input parameters.

Table 1
Comparison results.

		x_i follows $N(1, 0.5^2)$		x_i follows $U(0, 1)$		No. of
		Value	Error	Value	Error	Simulations
True Output Mean		**23.8047**	*	**2.9306**	*	*
M-PCM		23.8047	0	2.9306	0	64
M-PCM–OFFD	Cross-product power $\tau = 3$	23.8047	0	2.9306	0	64
	Cross-product power $\tau = 2$	23.8047	0	2.9306	0	32
	Cross-product power $\tau = 1$	23.8047	0	2.9306	0	8
SRSM	PCE of order 3	NaN	NaN	NaN	NaN	168
	PCE of order 2	24.2423	0.4376	3.0972	0.1666	56
	PCE of order 1	15.3269	8.4778	2.5570	0.3736	13
PCE-MCR	Formula IV, PCE of order 3	23.5210	0.2867	2.8771	0.0535	137
	Formula I, PCE of order 2	23.8047	0	3.1500	0.2194	44
PCE-SG (Gaussian)	Level $n = 4$, PCE of order 1,2,3	23.8047	0	2.9306	0	389
	Level $n = 3$, PCE of order 1,2,3	23.7813	0.0234	2.9301	0.0005	85
	Level $n = 2$, PCE of order 1,2,3	22.6250	1.1797	2.8646	0.0660	13
PCE-SG (KP)	Level $n = 4$, PCE of order 1,2,3	23.8047	0	2.9306	0	257
	Level $n = 3$, PCE of order 1,2,3	23.7813	0.0234	2.9301	0.0005	73
	Level $n = 2$, PCE of order 1,2,3	22.6250	1.1797	2.8646	0.0660	13
UDR	4 points for each parameter	22.6250	1.1797	2.8646	0.066	25
	3 points for each parameter	22.6250	1.1797	2.8646	0.066	19
	2 points for each parameter	22.6250	1.1797	2.8646	0.066	13
Coherence-Optimal Sampling	PCE of order 3	23.7273	0.0774	2.9319	0.0013	168
	PCE of order 2	25.3835	1.5788	2.8905	0.0401	56
	PCE of order 1	31.9276	8.1229	3.0750	0.1444	13

4.3. Comparative studies

In this section, we compare the performance of M-PCM–OFFD with some existing approaches reviewed in Section 1, including M-PCM, SRSM [32], PCE-MCR [63], PCE-SG [29,40,46], UDR [37], and coherence-optimal sampling [14,25,26].

Consider a 6-parameter system of the mapping $g(x_1, x_2, \ldots x_6) = 1.5 + x_5 - 0.5x_1^2 - 2x_6^3 + 5x_1^3x_2 + 4x_2^2x_4^3 + 3x_5^2x_6^2 + 2x_3x_5^2 + 0.5x_1^2x_2^2x_5^3 - 0.5x_2x_3^2x_4$, where all uncertain input parameters x_i follow either the normal distribution $N(1, 0.5^2)$ or the uniform distribution $U(0, 1)$. We aim to accurately estimate the output mean $E[g(x_1, x_2, \ldots x_6)]$ by running only a small number of simulations. Note that the original system mapping $g(x_1, x_2, \ldots x_6)$ is treated as a black box, with only inputs and outputs accessible. Table 1 summarizes the estimation results obtained using each uncertainty evaluation approach. The results are also compared to the true output mean $E[g(x_1, x_2, \ldots x_6)]$ calculated through integration using the explicit expression of $g(x_1, x_2, \ldots x_6)$.

M-PCM accurately estimates the output mean with $n_i = 2$ for $i \in \{1, 2, \ldots 6\}$, as proved in Theorem 1. The number of simulations is $2^6 = 64$. M-PCM–OFFD also estimates the output mean accurately by setting the maximal cross-product power $\tau = 3$. In this case, M-PCM–OFFD is equivalent to M-PCM as $\tau = 3 > \left\lceil \frac{6}{2} \right\rceil - 1$ (see Step 1 of Algorithm 3) and no further reduction can be achieved by OFFDs. If we set $\tau = 2$, a 2^{6-1}_{VI} OFFD can be applied to reduce the size of simulation set to 32. Furthermore, if we set $\tau = 1$, a 2^{6-3}_{V} OFFD can be used to further reduce the size of simulation set to 8. Of interest, in these two latter cases, despite that τ is smaller than what is needed to guarantee estimation correctness, the simulations still show excellent output-mean estimation performance. The special structure of OFFD reflected in the balance and orthogonality properties tends to minimize the impact of high-order cross-terms. Due to space limitation, we will investigate the performance bounds of M-PCM–OFFD for nonlinear, high-order systems, and high-order cross-terms in the future. Here τ can be determined through an iterative procedure. A larger value of τ results in better estimation accuracy but more simulation points. In cases when the computational cost is of little concern, τ can be determined by gradually increasing its value (starting from $\tau = 1$) until the estimated output

mean converges. The use of M-PCM–OFFD in the application of air traffic flow management and its capability in facilitating stochastic optimal control can be found in [65,66,71].

For PCE based approaches including SRSM, PCE-MCR, PCE-SG and coherence-optimal sampling, transformations are sometimes needed to first represent uncertain input parameters as standardized random variables. For instance, for SRSM and PCE-MCR that adopt PCE with variables following standard normal distributions, the uncertain input parameters can be converted to standard normal random variables α_i of the distribution $N(0, 1)$ using the following operations [32]:

$$x_i = \begin{cases} 1 + 0.5\alpha_i & \text{if } x_i \text{ follows } N(1, 0.5^2) \\ \frac{1}{2} + \frac{1}{2}\text{erf}\left(\frac{\alpha_i}{\sqrt{2}}\right) & \text{if } x_i \text{ follows } U(0, 1) \end{cases} \tag{14}$$

where $\text{erf}(x)$ is defined as $\text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt$. Then PCEs of order 3 have cross-terms containing up to 3 parameters to approximate the original mapping. We also tried lower-order PCEs for the comparison purpose. In our study of SRSM, a set of simulation points with size twice the number of PCE coefficients is selected, if enough points are available. The simulations show that the performance of SRSM decays significantly with the decrease of PCE order. When the applied PCE has an order of 3, SRSM often fails to calculate the output mean, as it randomly selects simulation points when no more points can be further eliminated after filtering.

For PCE-MCR, different formulas of MCR lead to different numbers of simulations. We here follow the simulation studies conducted in [63], and evaluate two formulas of MCR: (1) Formula I for PCE of order 2, and (2) Formula IV for PCE of order 3. The simulations imply that PCE-MCR has smaller estimation error compared to SRSM, but still cannot achieve accurate estimation of the output mean.

The PCE-SG provides a systematic procedure to generate a set of quadrature points to estimate each mapping coefficient. As the output mean of a PCE mapping equals the coefficient associated with the constant term, the accuracy level of the applied SG determines the accuracy of output mean estimation, rather than the order of PCE. Note that an SG of accuracy level n can correctly estimate the integral of a polynomial with the total order $2n - 1$. In this example, as the total order of the system mapping is 7, the SG of accuracy level $n = 4$ is capable of correctly estimating the output mean. In the simulation, we use PCE with Hermite polynomials and apply SG with Gauss–Hermite quadrature rules when uncertain input parameters x_i follow normal distributions. We use PCE with Legendre polynomials and apply the SG with the Gauss–Legendre quadrature rules when x_i follow uniform distributions [29]. Other than the standard Gaussian rules based SG, we also evaluate the Kronrod–Patternson (KP) rules based SG, which generates a nested set of points and is more efficient in high dimensions [10,29,46]. The simulations show that both Gaussian rules based and KP rules based SG have good estimation performance, but require many more simulation points compared to M-PCM–OFFD. The estimation performance decays with the decrease of the accuracy level n. Furthermore, we vary the order of PCE and verify the conclusion that the order of PCE does not have an impact on the estimation accuracy.

For UDR, the idea to calculate output mean is to obtain the output mean of each univariate function constructed for each uncertain input parameter. For fair comparison with M-PCM–OFFD, we here follow the procedures of Algorithm 1 (with $m = 1$) to select simulation points, construct the univariate functions, and calculate the output means. Note that n simulation points can accurately approximate the output mean of a univariate function of order up to $2n - 1$. Therefore, in this example, 2 simulation points are enough to accurately approximate a univariate function with the highest order up to 3. For verification, we also tested the performance of UDR when more simulation points are used to construct higher-order univariate functions. The simulations show that this approach does not perform well no matter how many points are used, as it ignores all cross-terms. In addition, the constant output means estimated for all three cases verify our statement on the number of simulation points that are sufficient for the output mean estimation.

For coherence-optimal sampling, we use the Matlab code available at www.github.com/CU-UQ to generate samples. The number of samples can be customized. We here adopt the same settings as the SRSM, as we found the estimation performance of the coherence-optimal sampling is not satisfactory when the number of samples is equal or slightly larger than the number of PCE coefficients. The simulations show that the performance of coherence-optimal sampling also degrades with the decrease of PCE order, but it outperforms SRSM. Besides, by adopting higher-order PCEs and using more samples, the output mean estimated by the coherence-optimal sampling is expected to converge to the accurate value.

Overall, among all these uncertainty evaluation approaches, M-PCM–OFFD performs the best in estimating the output mean in this example. It also requires fewer simulation points than all other approaches. Compared to PCE-based approaches, the UDR generally requires fewer simulations, but it does not perform well when cross-terms are not negligible. We also compare estimation errors for methods of similar number of simulations. In particular, we select M-PCM–OFFD of $\tau = 1$, SRSM and coherence-optimal sampling when the PCE is of order 1, UDR with 2 points selected for each parameter, and SG with accuracy level $n = 2$. As we can see in Table 1, M-PCM–OFFD performs the best, while SRSM produces the largest error.

5. Conclusions

The real-time decision-making of large-scale complex systems requires an effective method to evaluate statistical performance for systems of high-dimensional uncertainties. In this paper, we integrate M-PCM with OFFD, and show that the integrated method maintains the good estimation performance of M-PCM, while significantly reduces the number of simulations. In particular, we found that under the conditions that the original system mapping of m parameters has the degree of each parameter up to 3 and the maximal cross-product power satisfies $\tau \leq \lceil \frac{m}{2} \rceil - 1$, the reduced-order polynomial mapping constructed using the M-PCM–OFFD method has the following features: (1) it precisely estimates the output mean of original system mapping; (2) it reduces the number of simulations from 2^m to the range of $[2^{\lceil \log_2(m+1) \rceil}, 2^{m-1}]$, breaking the curse of dimensionality; (3) it is the most robust to numerical truncation errors of simulators among all subsets of the same size in the M-PCM simulation set to meet parameter resolution requirements; and (4) it correctly estimates main and significant interaction effects. The theoretical analysis developed in this paper on the estimation correctness, computational scalability, and robustness to numerical truncation errors of simulators demonstrate appealing properties of M-PCM–OFFD for its practical use in developing fast decision-making solutions for large-scale system applications. The development in this paper also provides new insights into the optimality of OFFDs, and gives rise to its broad new usage for real-time uncertainty evaluation applications. In our future work, we will generalize the degrees of uncertain input parameters by exploring multi-level OFFDs and exploit parameter correlations to further reduce the number of simulations [1,7,9,38,53]. We will also investigate performance bounds for general systems of nonlinear mappings and high-order cross-terms.

Acknowledgments

Wan and Xie would like to thank the National Institute of Standards and Technology (NIST), United States for the support of this work, and also National Science Foundation, United States under Grants CNS1714826, CNS1714519, ECCS 1839707, and ECCS 1839804 for the partial support. The authors would also like to acknowledge Yi Zhou at Southwest Airlines for some preliminary studies and Christopher E. Dabrowski at NIST for valuable review comments.

Appendix

A.1. Proof of Lemma 1

According to Step 1 in Algorithm 2, any $m - \gamma$ factors are determined according to the full factorial design. The values of all other factors are determined using a subset of these $m - \gamma$ factors according to the generators. As the resolution R equals the shortest length of all these generators, $m - \gamma \geq R - 1$ holds.

A.2. Proof of Lemma 2

According to the proofs for Theorems 1 and 2 in [71], M-PCM recursively reduces the degree of each input parameter to produce a reduced-order mapping of the same output mean. As this procedure does not introduce new parameters to each cross-term, the numbers of parameters in all cross-terms in the reduced-order mapping do not increase.

A.3. Proof of Lemma 3

We first prove that an OFFD design can further reduce the number of simulations if $1 \leq \tau \leq \lceil \frac{m}{2} \rceil - 1$ and $m > 2$. Step 2 of Algorithm 3 suggests that the number of M-PCM–OFFD coefficients for a mapping of maximal cross-product power τ is $\sum_{i=0}^{\tau} \binom{m}{i}$. As OFFDs reduce the number of simulations at least by half, the number of coefficients $\sum_{i=0}^{\tau} \binom{m}{i}$ in $g^*(x_1, x_2, \ldots x_m)$ must be less than or equal to 2^{m-1} to ensure correct estimation. Note that $2^{m-1} = \sum_{i=0}^{\frac{m-1}{2}} \binom{m}{i}$ when m is odd and $\sum_{i=0}^{\frac{m}{2}-1} \binom{m}{i} < 2^{m-1} < \sum_{i=0}^{\frac{m}{2}} \binom{m}{i}$ when m is even. The maximum of τ thus satisfies $\max \tau = \begin{cases} \frac{m-1}{2} & \text{if } m \bmod 2 \equiv 1 \\ \frac{m}{2} - 1 & \text{if } m \bmod 2 \equiv 0 \end{cases} = \lceil \frac{m}{2} \rceil - 1$. As the maximal cross-product power τ is an integer greater than or equal to 1, it needs to be in the range of $1 \leq \tau \leq \lceil \frac{m}{2} \rceil - 1$.

We next prove that the maximal reduction of simulations can be achieved using the $2^{m-\gamma_{\max}}$ OFFD, where $\gamma_{\max} = m - \lceil \log_2 \left(\sum_{i=0}^{\tau} \binom{m}{i} \right) \rceil$. This is straightforward, as the number of simulations $2^{m-\gamma}$ must be larger than or equal to the number of M-PCM–OFFD coefficients $\sum_{i=0}^{\tau} \binom{m}{i}$ to ensure the correct estimation of mapping coefficients in $g^+(x_1, x_2, \ldots x_m)$ or $g^*(x_1, x_2, \ldots x_m)$. Therefore, $\gamma_{\max} = \max \left\{ \gamma \mid 2^{m-\gamma} \geq \sum_{i=0}^{\tau} \binom{m}{i} \right\} = m - \lceil \log_2 \left(\sum_{i=0}^{\tau} \binom{m}{i} \right) \rceil$.

A.4. Proof of Lemma 4

The first column of L'_{offd} is $\mathbf{1}$, and hence $\mathbf{v}_1 = \mathbf{1}$. For $1 < i \leq m + 1$, $\mathbf{v}_i$ contains the coded values of parameter x_{i-1}, $x_{c_{i-1}}$. According to Step 5 of Algorithm 3, all rows in L'_{offd} that are not in the OFFD design table are eliminated. It is thus clear that $\mathbf{v}_i = \mathbf{d}_{i-1}$ for $1 < i \leq m + 1$. The construction of L'_{offd} in Algorithm 3 also indicates that its i th column, where $i > m + 1$, is an element-wise vector multiplication of a subset of its first $m + 1$ columns. $\mathbf{v}_i = \prod_{k \in S_i} \mathbf{v}_k$ is thus derived.

A.5. Proof of Lemma 5

According to Lemma 4 and the orthogonality and balance properties of OFFD design table stated in Section 2.2, it is straightforward that $\mathbf{v}_i \cdot \mathbf{v}_j = 0$ when $i, j \in \{1, 2, \ldots m, m+1\}$. For arbitrary i and j where $i \neq j$, $\mathbf{v}_i \cdot \mathbf{v}_j = \mathbf{v}_i \circ \mathbf{v}_j \cdot \mathbf{v}_1 = \prod_{k \in S_i} \mathbf{v}_{k+1} \circ \prod_{k \in S_j} \mathbf{v}_{k+1} \cdot \mathbf{v}_1 = \prod_{k \in S} \mathbf{v}_{k+1} \cdot \mathbf{v}_1$, where $S = S_i \cup S_j$. According to Lemma 3, the sizes of S_i and S_j produced from V for the reduced-order mapping are both less than τ, and hence the size of S must be smaller than or equal to 2τ. Note that in a $2^{m-\gamma}_R$ OFFD, any $m - \gamma$ factors form a full factorial design (Step 1 of Algorithm 2). Hence, $\prod_{k \in S'} \mathbf{v}_{k+1} \cdot \mathbf{v}_1 = 0$ for an arbitrary S' whose size is less than or equal to $m - \gamma$, according to the orthogonality property of full factorial design discussed in Section 2.2. Therefore, to prove this lemma, we only need to show that $2\tau \leq m - \gamma$. According to Lemma 2, $m - \gamma \geq R - 1$. When $R \geq 2\tau + 1$ (Step 2 of Algorithm 3), $m - \gamma \geq R - 1 \geq 2\tau$ holds.

A.6. Proof of Lemma 6

Note that for each parameter x_k, its PCM points can be expressed as $x_{k(1)} = \frac{1}{2}\left(x_{k(1)} + x_{k(2)}\right) - \frac{1}{2}\left(x_{k(2)} - x_{k(1)}\right)$, and $x_{k(2)} = \frac{1}{2}\left(x_{k(1)} + x_{k(2)}\right) + \frac{1}{2}\left(x_{k(2)} - x_{k(1)}\right)$. Define $\frac{1}{2}\left(x_{k(1)} + x_{k(2)}\right) = x_{k,0}$ and $y_k = x_k - x_{k,0}$. We have $x_{k(1)} = x_{k,0} - \frac{1}{2}\Delta x_k$, $x_{k(2)} = x_{k,0} + \frac{1}{2}\Delta x_k$, and $y_{k(1)} = -\frac{1}{2}\Delta x_k$, $y_{k(2)} = \frac{1}{2}\Delta x_k$. The element in the r th row and i th column of matrix L'_{offd} can then be represented as $\prod_{k \in S_i} x_{k(r_k)} = \prod_{k \in S_i} y_{k(r_k)} + x_{k,0} = \prod_{k \in S_i} y_{k(r_k)} + \sum_{S_j \subset S_i} c_j \prod_{k \in S_j} y_{k(r_k)}$, where $r_k \in \{1, 2\}$ and c_j are some constants. Construct a matrix $K'_{\text{offd}} \in R^{l_{\text{offd}} \times l}$ from L'_{offd}, with $x_{k(1)}$ and $x_{k(2)}$ in L'_{offd} replaced by $y_{k(1)} = -\frac{1}{2}\Delta x_k$ and $y_{k(2)} = \frac{1}{2}\Delta x_k$, respectively. The expression above about $\prod_{k \in S_i} x_{k(r_k)}$ suggests that there exists an upper triangular connection matrix $U_1 \in R^{l \times l}$ of unity diagonal terms (and off-diagonal elements given by the c_j constants above) that satisfies $L'_{\text{offd}} = K'_{\text{offd}} U_1$.

According to the definitions of the K'_{offd} and V matrices, we have $K'_{\text{offd}} = V U_2$, where $U_2 \in R^{l \times l}$ is a diagonal

matrix with the i th diagonal element equal to 1 if $i = 1$ and $\frac{1}{2^{\xi_i}} \prod_{k \in S_i} \Delta x_k$ otherwise. We then have $L'_{\text{offd}} = VU$, where $U = U_1 U_2$ is an upper triangular matrix with the same diagonal elements as matrix U_2. Matrix V has orthogonal columns according to Lemma 5, and it satisfies $\frac{1}{l_{\text{offd}}} V^T V = I$. Therefore, both matrices V and U have full column ranks, and hence matrix L'_{offd} also has full column rank.

A.7. Proof of Theorem 2

Theorem 1 states that the reduced-order M-PCM mapping $g^+(x_1, x_2, \ldots x_m)$ correctly estimates the output mean of $g(x_1, x_2, \ldots x_m)$, with a reduction of simulations from 2^{2m} to 2^m. Lemma 2 states that M-PCM mapping $g^+(x_1, x_2, \ldots x_m)$ has the same maximal cross-product power τ as the M-PCM–OFFD mapping $g^*(x_1, x_2, \ldots x_m)$. Lemma 3 states that there exists an OFFD to further reduce the number of simulations, when $1 \le \tau \le \lceil \frac{m}{2} \rceil - 1$ and $m > 2$ (Step 1 of Algorithm 3). Lemma 6 states that when $R \ge 2\tau + 1$, the $2^{m-\gamma^*}_R$ OFFD results in a full-column-rank matrix L'_{offd}, and hence the calculation in Step 5 of Algorithm 3 is feasible. We note that when $1 \le \tau \le \lceil \frac{m}{2} \rceil - 1$ and $m > 2$, such a $2^{m-\gamma^*}_R$ OFFD always exists, where γ^* satisfies $1 \le \gamma^* \le m - \lceil \log_2 (\sum_{i=0}^{\tau} \binom{m}{i}) \rceil$, and $R \ge 2\tau + 1$. In particular, the 2^{m-1}_m OFFD produced using the generators $I = \pm x_{c_1} x_{c_2} \cdots x_{c_m}$ always exists, and meets the conditions, as $\gamma = 1$ and $R = m \ge 2\lceil \frac{m}{2} \rceil - 1 \ge 2\tau + 1$. In all, the M-PCM–OFFD mapping $g^*(x_1, x_2, \ldots x_m)$ is the same as the M-PCM mapping $g^+(x_1, x_2, \ldots x_m)$, despite the reduction of simulation points, as all the M-PCM points are precisely on the mapping g^*. The correct output mean estimation naturally follows. Furthermore, according to Lemma 3–6, the number of simulations is reduced to $2^{m-\gamma^*}$, where $\gamma^* = \max\{\gamma \mid 1 \le \gamma \le m - \lceil \log_2 (\sum_{i=0}^{\tau} \binom{m}{i}) \rceil$, *and* $2^{m-\gamma}_R$ *OFFD exists*, *with* $R \ge 2\tau + 1\}$. The lower bound of $2^{m-\gamma^*}$ is achieved when $\tau = 1$, and the upper bound holds as an OFFD at least halves the number of simulations.

A.8. Proof of Lemma 7

According to Lemma 6, L'_{offd} is of full column rank and $L'_{\text{offd}} = VU$. Now we find the minimum $\|e\|_F$ to make $L'_{\text{offd}} + e$ lose rank, according to the definition of full-column-rank margin in (11). We use $e_{x_{i(r_i)}}$ to represent the perturbation to $x_{i(r_i)}$ and $\hat{x}_{i(r_i)} = x_{i(r_i)} + e_{x_{i(r_i)}}$ to represent the corrupted parameter value, where $i = \{1, 2, \ldots m\}$, and $r_i = \{1, 2\}$. Similar to L'_{offd}, $L'_{\text{offd}} + e$ can also be expressed as $L'_{\text{offd}} + e = V\hat{U}$, where $\hat{U}$ is an upper triangular matrix with the determinant $det(\hat{U}) = \prod_{i=2}^{l} (\frac{1}{2^{\xi_i}} \prod_{k \in S_i} \Delta \hat{x}_k)$, where $\Delta \hat{x}_k = \hat{x}_{k(2)} - \hat{x}_{k(1)}$. Clearly, the rank of L'_{offd} is solely determined by $\hat{U}$. Therefore, $L'_{\text{offd}} + e$ loses rank if and only if there exists an $i \in \{1, 2, \ldots m\}$ such that $\Delta \hat{x}_i = 0$. In the case of $\Delta \hat{x}_1 = 0$, we have $\Delta \hat{x}_1 = \hat{x}_{1(2)} - \hat{x}_{1(1)} = (x_{1(2)} + e_{x_{1(2)}}) - (x_{1(1)} + e_{x_{1(1)}}) = 0$ and therefore $e_{x_{1(1)}} = e_{x_{1(2)}} + x_{1(2)} - x_{1(1)} = e_{x_{1(2)}} + \Delta x_1$. Consequently,

$$\|e\|_F = \sqrt{\frac{l_{\text{offd}}}{2} \left(e^2_{x_{1(1)}} + e^2_{x_{1(2)}} + e^2_{x_{2(1)}} + e^2_{x_{2(2)}} + \cdots + e^2_{x_{m(1)}} + e^2_{x_{m(2)}} \right)}$$
$$\ge \sqrt{\frac{l_{\text{offd}}}{2} \left(e^2_{x_{1(1)}} + e^2_{x_{1(2)}} \right)} = \sqrt{\frac{l_{\text{offd}}}{2} \left[(e_{x_{1(2)}} + \Delta x_1)^2 + e^2_{x_{1(2)}} \right]} \ge \sqrt{\frac{l_{\text{offd}}}{2}} \Delta x_1$$

The equality holds when $e_{x_{1(1)}} = \frac{1}{2}\Delta x_1$, $e_{x_{1(2)}} = -\frac{1}{2}\Delta x_1$, and $e_{x_{j(1)}} = e_{x_{j(2)}} = 0$ for all $j = 1$. Similarly, we obtain $\|e\|_F \ge \sqrt{\frac{l_{\text{offd}}}{2}} \Delta x_i$, $i \in \{2, 3, \ldots m\}$. As such, $M(L'_{\text{offd}}) = \sqrt{\frac{l_{\text{offd}}}{2}} \min\{\Delta x_1, \Delta x_2, \ldots \Delta x_m\}$, and the minimum $\sqrt{\frac{l_{\text{offd}}}{2}} \Delta x_i$ is achieved when $\Delta x_i \le \Delta x_j$ for all $j = i$, $e_{x_{i(1)}} = \frac{1}{2}\Delta x_i$, $e_{x_{i(2)}} = -\frac{1}{2}\Delta x_i$, and $e_{x_{j(1)}} = e_{x_{j(2)}} = 0$ for all $j = i$.

A.9. Proof of Theorem 3

Before showing the proof of Theorem 3, we first present a lemma, which will be used to prove this theorem.

Lemma 9. *Consider an original system mapping* $g(x_1, x_2, \ldots x_m)$, (7), *with* $m > 2$ *and* $\tau = 1$. *Matrix* L' *constructed from an arbitrary subset of size* $2^{\lceil log_2(m+1) \rceil}$ *in the M-PCM simulation set can be transformed to an upper triangular matrix of the same column rank, where the first diagonal entry is* 1, *and the* $(k+1)$*th diagonal entry is an integer multiple of* Δx_k.

Proof. Denote the entry of matrix L' at the i th row and j th column as $L'_{i,j}$. In cases when $\tau = 1$, matrix L' consists of $m+1$ columns arranged in the following order: $\{1, x_1, x_2, \ldots x_m\}$ (see Remark 2 of Algorithm 3 for the ordering criteria). As the $(k+1)$th column of matrix L' contains only two possible values, $x_{k(1)}$ and $x_{k(2)}$, $k \in \{1, 2, \ldots m\}$, matrix L' can be transformed to an upper triangular form, with diagonal entries being integer multiples of Δx_k, through a sequence of *elementary row operations*: for $i = 1$ to $m-1$ do the following; (1) make $L'_{j,i} = 0$ by subtracting the j th row with a multiple of the i th row, for $j = i+1, \ldots m$; (2) if $L'_{i+1,i+1} = 0$, find a non-zero element $L'_{k,i+1}$, $k = i+2, \ldots m$, and switch row k with row $i+1$. Note that the elementary row operations do not change the column rank of matrix L' [56].

Now let us find the minimum $\|e\|_F$ to make $L'+e$ lose rank. The case that L' is not of full column rank is trivial, as in this case the minimum $\|e\|_F = 0$ and $M(L') = 0$, where e is a null matrix. When L' is full rank, we can transform, through elementary row operations described in the proof of Lemma 9, $L'+e$ to an upper triangular matrix, in which the first diagonal entry is 1, and the $(k+1)$th diagonal entry is an integer multiple of $\Delta\hat{x}_k$, i.e., $\lambda_k \Delta\hat{x}_k$, where $\lambda_k \in Z$ and Z denotes the set of integers. Matrix L' being of full column rank implies that $\lambda_k \neq 0$, for all $k \in \{1, 2, \ldots m\}$. As such, $L'+e$ loses column rank if and only if $\Delta\hat{x}_k = 0$.

Procedures similar to those used to calculate $M(L'_{\text{offd}})$ in the proof of Lemma 7 lead to

$$\begin{aligned} M(L') &= \min\left\{\sqrt{\left[c_1 e^2_{x_{1(1)}} + (l_{\text{offd}} - c_1)e^2_{x_{1(2)}}\right] + \cdots + \left[c_m e^2_{x_{m(1)}} + (l_{\text{offd}} - c_m)e^2_{x_{m(2)}}\right]}\right\} \\ &= \min\left\{\sqrt{\frac{c_1(l_{\text{offd}} - c_1)}{l_{\text{offd}}}}\Delta x_1, \ldots, \sqrt{\frac{c_m(l_{\text{offd}} - c_m)}{l_{\text{offd}}}}\Delta x_m\right\} \end{aligned} \tag{15}$$

where c_i is the number of $x_{i(1)}$ in the $(i+1)$th column of L'. The minimum at $\sqrt{\frac{c_i(l_{\text{offd}} - c_i)}{l_{\text{offd}}}}\Delta x_i$ is achieved, when $e_{x_{i(1)}} = \frac{(l_{\text{offd}} - c_i)\Delta x_i}{l_{\text{offd}}}$, $e_{x_{i(2)}} = -\frac{c_i \Delta x_i}{l_{\text{offd}}}$, and for all $j \neq i$, $j \in \{1, 2, \ldots m\}$, we have $\sqrt{\frac{c_i(l_{\text{offd}} - c_i)}{l_{\text{offd}}}}\Delta x_i \leq \sqrt{\frac{c_j(l_{\text{offd}} - c_j)}{l_{\text{offd}}}}\Delta x_j$, and $e_{x_{j(1)}} = e_{x_{j(2)}} = 0$.

Since $\sqrt{\frac{c_i(l_{\text{offd}} - c_i)}{l_{\text{offd}}}} = \sqrt{-\frac{1}{l_{\text{offd}}}(c_i - \frac{l_{\text{offd}}}{2})^2 + \frac{l_{\text{offd}}}{4}} \leq \sqrt{\frac{l_{\text{offd}}}{2}}$, we have $\sqrt{\frac{c_i(l_{\text{offd}} - c_i)}{l_{\text{offd}}}}\Delta x_i \leq \sqrt{\frac{l_{\text{offd}}}{2}}\Delta x_i$. (15) can then be further simplified to $M(L') \leq \sqrt{\frac{l_{\text{offd}}}{2}} \min\{\Delta x_1, \Delta x_2, \ldots \Delta x_m\} = M(L'_{\text{offd}})$. The equality is achieved by an OFFD.

A.10. Proof of Lemma 8

To prove Lemma 8, we first construct $\|e\|_F$ and then find its minimum to make L'_{offd} lose rank. Note that the $L'_{\text{offd}}+e$ matrix has the same structure as that of L'_{offd}, with its i th column $\prod_{k\in S_i}\hat{x}_k = \prod_{k\in S_i}(x_k + e_{x_k})$. Simple algebra leads to

$$\begin{aligned} \|e\|_F = \Bigg\{ &\frac{l_{\text{offd}}}{2}\sum_{k=1}^{m}\sum_{r_k=1}^{2} e^2_{x_{k(r_k)}} + \frac{l_{\text{offd}}}{2^2}\sum_{r_1=1}^{2}\sum_{r_2=1}^{2}\left[\left(e_{x_{1(r_1)}} + x_{1(r_1)}\right)\left(e_{x_{2(r_2)}} + x_{2(r_2)}\right)\right. \\ &\left. - x_{1(r_1)}x_{2(r_2)}\right]^2 + \cdots + \frac{l_{\text{offd}}}{2^\tau}\left\{\sum_{r_1=1}^{2}\cdots\sum_{r_\tau=1}^{2}\left[\prod_{k=1}^{\tau}\left(e_{x_{k(r_k)}} + x_{k(r_k)}\right) - \prod_{k=1}^{\tau} x_{k(r_k)}\right]^2\right\} + \cdots \end{aligned} \tag{16}$$

Similar to the proof of Lemma 7, $L'_{\text{offd}}+e$ loses rank if and only if at least one of $\Delta\hat{x}_i = 0$, $i \in A$. We assume WLOG $\Delta\hat{x}_1 = 0$, and thus $e_{x_{1(1)}} + x_{1(1)} = e_{x_{1(2)}} + x_{1(2)}$, and $e_{x_{1(1)}} = e_{x_{1(2)}} + \Delta x_1$.

In this case, to find the minimum $\|e\|_F$ for L'_{offd} to lose rank, we show that the minimum of each error summation term corresponding to each column of $L'+e$ in (16) and hence the minimum of $\|e\|_F$ is achieved when $e_{x_{1(1)}} = \frac{1}{2}\Delta x_1$, $e_{x_{1(2)}} = -\frac{1}{2}\Delta x_1$, and $e_{x_{j(1)}} = e_{x_{j(2)}} = 0$ for all $j \neq 1$, $j \in A$.

First, for columns of $L'+e$ that contain only one input parameter, we have $\sum_{k=1}^{m}\sum_{r_k=1}^{2} e^2_{x_{k(r_k)}} \geq \sum_{r_1=1}^{2} e^2_{x_{1(r_1)}} \geq \frac{\Delta x_1^2}{2}$. The minimum value is achieved when $e_{x_{1(1)}} = \frac{1}{2}\Delta x_1$, $e_{x_{1(2)}} = -\frac{1}{2}\Delta x_1$, and $e_{x_{j(1)}} = e_{x_{j(2)}} = 0$ for all $j \neq 1$, $j \in A$.

Next, for columns that contain two input parameters ($\tau \geq 2$), we consider two cases. In the first case, the column does not contain the input parameter x_1. The minimum of corresponding error summation term is achieved when no error is introduced. That is,

$$\sum_{r_i=1}^{2}\sum_{r_j=1}^{2}\left[(e_{x_{i(r_i)}}+x_{i(r_i)})(e_{x_{j(r_j)}}+x_{j(r_j)})-x_{i(r_i)}x_{j(r_j)}\right]^2\geq 0$$

where $i, j = 1$, and the minimum is achieved when $e_{x_{i(1)}} = e_{x_{i(2)}} = e_{x_{j(1)}} = e_{x_{j(2)}} = 0$. For columns that contain input parameter x_1, we have

$$\begin{aligned}
&\sum_{r_1=1}^{2}\sum_{r_i=1}^{2}\left[\left(e_{x_{1(r_1)}}+x_{1(r_1)}\right)\left(e_{x_{i(r_i)}}+x_{i(r_i)}\right)-x_{1(r_1)}x_{i(r_i)}\right]^2\left[\left(e_{x_{1(1)}}+x_{1(1)}\right)\left(e_{x_{i(1)}}+x_{i(1)}\right)-x_{1(1)}x_{i(1)}\right]^2\\
&+\left[\left(e_{x_{1(2)}}+x_{1(2)}\right)\left(e_{x_{i(1)}}+x_{i(1)}\right)-x_{1(2)}x_{i(1)}\right]^2+\left[\left(e_{x_{1(1)}}+x_{1(1)}\right)\left(e_{x_{i(2)}}+x_{i(2)}\right)-x_{1(1)}x_{i(2)}\right]^2\\
&+\left[\left(e_{x_{1(2)}}+x_{1(2)}\right)\left(e_{x_{i(2)}}+x_{i(2)}\right)-x_{1(2)}x_{i(2)}\right]^2\\
&=2\left[\left(e_{x_{1(1)}}+x_{1(1)}\right)\left(e_{x_{i(1)}}+x_{i(1)}\right)-\frac{1}{2}\left(x_{1(1)}+x_{1(2)}\right)x_{i(1)}\right]^2+\frac{1}{2}x_{i(1)}^2\Delta x_1^2\\
&+2\left[\left(e_{x_{1(1)}}+x_{1(1)}\right)\left(e_{x_{i(2)}}+x_{i(2)}\right)-\frac{1}{2}\left(x_{1(1)}+x_{1(2)}\right)x_{i(2)}\right]^2+\frac{1}{2}x_{i(2)}^2\Delta x_1^2\\
&\geq\frac{1}{2}\Delta x_1^2\left(x_{i(1)}^2+x_{i(2)}^2\right)=\frac{1}{2}\Delta x_1^2\prod_{a\in A_{1,1,i-1}}\left(x_{a(1)}^2+x_{a(2)}^2\right)
\end{aligned}$$

where $A_{1,1,i-1} = \{i\}$. The minimum value can be achieved when $e_{x_{1(1)}} = \frac{1}{2}\Delta x_1$, $e_{x_{1(2)}} = -\frac{1}{2}\Delta x_1$, and $e_{x_{i(1)}} = e_{x_{i(2)}} = 0$, where $i = 1$, $i \in A$.

For terms that contain s number of input parameters, $s \in \{3, 4, \ldots, (\tau \geq 3)\}$, the same conclusion can be obtained by following similar procedures. In particular, for columns that do not contain x_1, it is clear that the minimum of error summation is 0, which is achieved when no error is introduced. For columns that contain x_1, the minimum of error summation is achieved when input parameters except x_1 in this column is free of error. We here show, WLOG, the calculation for the column that contains input parameters $\{x_1, x_2, \ldots x_s\}$,

$$\begin{aligned}
&\sum_{r_1=1}^{2}\cdots\sum_{r_s=1}^{2}\left[\prod_{k=1}^{s}\left(e_{x_{k(r_k)}}+x_{k(r_k)}\right)-\prod_{k=1}^{s}x_{k(r_k)}\right]^2\\
&=2\left(e_{x_{1(1)}}+x_{1(1)}\right)^2\left(e_{x_{2(1)}}+x_{2(1)}\right)^2\prod_{k=3}^{s}\left[\sum_{r_k=1}^{2}\left(e_{x_{k(r_k)}}+x_{k(r_k)}\right)^2\right]\\
&-2\left(e_{x_{1(1)}}+x_{1(1)}\right)\left(e_{x_{2(1)}}+x_{2(1)}\right)\left(x_{1(1)}+x_{1(2)}\right)x_{2(1)}\prod_{k=3}^{s}\left[\sum_{r_k=1}^{2}x_{k(r_k)}\left(e_{x_{k(r_k)}}+x_{k(r_k)}\right)\right]\\
&+\left(x_{1(1)}^2+x_{1(2)}^2\right)x_{2(1)}^2\prod_{k=3}^{s}\left(x_{k(1)}^2+x_{k(2)}^2\right)+2\left(e_{x_{1(1)}}+x_{1(1)}\right)^2\left(e_{x_{2(2)}}+x_{2(2)}\right)^2\prod_{k=3}^{s}\left[\sum_{r_k=1}^{2}\left(e_{x_{k(r_k)}}+x_{k(r_k)}\right)^2\right]\\
&-2\left(e_{x_{1(1)}}+x_{1(1)}\right)\left(e_{x_{2(2)}}+x_{2(2)}\right)\left(x_{1(1)}+x_{1(2)}\right)x_{2(2)}\prod_{k=3}^{s}\left[\sum_{r_k=1}^{2}x_{k(r_k)}\left(e_{x_{k(r_k)}}+x_{k(r_k)}\right)\right]
\end{aligned}$$

$$+\left(x_{1(1)}^2+x_{1(2)}^2\right)x_{2(2)}^2\prod_{k=3}^{s}\left(x_{k(1)}^2+x_{k(2)}^2\right)$$

$$=2\prod_{k=3}^{s}\left[\sum_{r_k=1}^{2}\left(e_{x_{k(r_k)}}+x_{k(r_k)}\right)_2\right]\left\{\left(e_{x_{1(1)}}+x_{1(1)}\right)\left(e_{x_{2(1)}}+x_{2(1)}\right)\right.$$

$$\left.-\frac{1}{2}x_{2(1)}\left(x_{1(1)}+x_{1(2)}\right)\frac{\prod_{k=3}^{s}\left[\sum_{r_k=1}^{2}x_{k(r_k)}\left(e_{x_{k(r_k)}}+x_{k(r_k)}\right)\right]}{\prod_{k=3}^{s}\left[\sum_{r_k=1}^{2}\left(e_{x_{k(r_k)}}+x_{k(r_k)}\right)_2\right]}\right\}^2-\frac{1}{2}x_{2(1)}^2\left(x_{1(1)}+x_{1(2)}\right)_2$$

$$\frac{\prod_{k=3}^{s}\left[\sum_{r_k=1}^{2}x_{k(r_k)}\left(e_{x_{k(r_k)}}+x_{k(r_k)}\right)\right]_2}{\prod_{k=3}^{s}\left[\sum_{r_k=1}^{2}\left(e_{x_{k(r_k)}}+x_{k(r_k)}\right)_2\right]}+\left(x_{1(1)}^2+x_{1(2)}^2\right)x_{2(1)}^2\prod_{k=3}^{s}\sum_{r_k=1}^{2}x_{k(r_k)}^2+\cdots$$

$$\geq\left(x_{2(1)}^2+x_{2(2)}^2\right)\left\{\left(x_{1(1)}^2+x_{1(2)}^2\right)\prod_{k=3}^{s}\sum_{r_k=1}^{2}x_{k(r_k)}^2-\frac{1}{2}\left(x_{1(1)}+x_{1(2)}\right)_2\frac{\prod_{k=3}^{s}\left[\sum_{r_k=1}^{2}x_{k(r_k)}\left(e_{x_{k(r_k)}}+x_{k(r_k)}\right)\right]_2}{\prod_{k=3}^{s}\left[\sum_{r_k=1}^{2}\left(e_{x_{k(r_k)}}+x_{k(r_k)}\right)_2\right]}\right\}$$

By using the Cauchy–Schwarz Inequality [57], we have

$$\left(x_{2(1)}^2+x_{2(2)}^2\right)\left\{\left(x_{1(1)}^2+x_{1(2)}^2\right)\prod_{k=3}^{s}\sum_{r_k=1}^{2}x_{k(r_k)}^2\right.$$

$$\left.-\frac{1}{2}\left(x_{1(1)}+x_{1(2)}\right)_2\frac{\prod_{k=3}^{s}\left[\sum_{r_k=1}^{2}x_{k(r_k)}\left(e_{x_{k(r_k)}}+x_{k(r_k)}\right)\right]_2}{\prod_{k=3}^{s}\left[\sum_{r_k=1}^{2}\left(e_{x_{k(r_k)}}+x_{k(r_k)}\right)_2\right]}\right\}$$

$$\geq\left(x_{2(1)}^2+x_{2(2)}^2\right)\left\{\left(x_{1(1)}^2+x_{1(2)}^2\right)\prod_{k=3}^{s}\sum_{r_k=1}^{2}x_{k(r_k)}^2-\frac{1}{2}\left(x_{1(1)}+x_{1(2)}\right)_2\right.$$

$$\left.\frac{\prod_{k=3}^{s}\left(\sum_{r_k=1}^{2}x_{k(r_k)}^2\right)\prod_{k=3}^{s}\left[\sum_{r_k=1}^{2}\left(e_{x_{k(r_k)}}+x_{k(r_k)}\right)_2\right]}{\prod_{k=3}^{s}\left[\sum_{r_k=1}^{2}\left(e_{x_{k(r_k)}}+x_{k(r_k)}\right)_2\right]}\right\}$$

$$=\left(x_{2(1)}^2+x_{2(2)}^2\right)\prod_{k=3}^{s}\sum_{r_k=1}^{2}x_{k(r_k)}^2\left[\left(x_{1(1)}^2+x_{1(2)}^2\right)-\frac{1}{2}\left(x_{1(1)}+x_{1(2)}\right)_2\right]$$

$$=\frac{1}{2}\Delta x_1^2\prod_{k=2}^{s}\left(x_{k(1)}^2+x_{k(2)}^2\right)=\frac{1}{2}\Delta x_1^2\prod_{a\in A_{1,s-1,1}}\left(x_{a(1)}^2+x_{a(2)}^2\right)$$

where $\mathcal{A}_{1,s-1,1} = \{2, 3, \ldots s\}$. The minimum value is achieved when $e_{x_{1(1)}} = \frac{1}{2}\Delta x_1$, $e_{x_{1(2)}} = -\frac{1}{2}\Delta x_1$, and $e_{x_{j(1)}} = e_{x_{j(2)}} = 0$ for all $j \in \{2, \ldots s\}$. Based on all above analyses, we see that the minimum value of each error summation term in (16) in the case of $\Delta\hat{x}_1 = 0$ is achieved at the same condition, i.e., $e_{x_{1(1)}} = \frac{1}{2}\Delta x_1$, $e_{x_{1(2)}} = -\frac{1}{2}\Delta x_1$, and $e_{x_{j(1)}} = e_{x_{j(2)}} = 0$ for all $j \neq 1, j \in \mathcal{A}$. Therefore, we have $\|e\|_F \geq \frac{\frac{1}{2}\Delta x_1}{\sqrt{1 + \sum_{i=1}^{\tau-1}\left[\frac{1}{2^i}\sum_{j=1}^{\binom{m-1}{i}}\prod_{a\in\mathcal{A}_{1,i,j}}\left(x_{a(1)}^2 + x_{a(2)}^2\right)\right]}}$, when $\Delta\hat{x}_1 = 0$. Similarly, we can obtain

$$\|e\|_F \geq \frac{\sqrt{l_{\text{offd}}}}{2}\Delta x_k \sqrt{1 + \sum_{i=1}^{\tau-1}\left[\frac{1}{2^i}\sum_{j=1}^{\binom{m-1}{i}}\prod_{a\in\mathcal{A}_{k,i,j}}\left(x_{a(1)}^2 + x_{a(2)}^2\right)\right]}$$

when $\Delta\hat{x}_k = 0$. The minimum is achieved when $e_{x_{k(1)}} = \frac{1}{2}\Delta x_k$, $e_{x_{k(2)}} = -\frac{1}{2}\Delta x_k$, and $e_{x_{j(1)}} = e_{x_{j(2)}} = 0$ for all $j \neq k, j \in \mathcal{A}$. The result naturally follows, similar to the proof of the first part for Lemma 7.

A.11. Proof of Theorem 4

To prove that $M(L') \leq M(L'_{\text{offd}})$, it suffices to show the existence of a perturbation matrix e which makes $L' + e$ lose rank and has a norm smaller than or equal to $M(L'_{\text{offd}})$. To find such a perturbation matrix, we first study the structure of L'. Note that matrix L' can be partitioned into two matrices $L' = \begin{bmatrix} L'_1 & L'_2 \end{bmatrix}$, where $L'_1 \in R^{l_{\text{offd}}\times(m+1)}$ contains the first $m+1$ columns, each of which has at most one input parameter. $L'_2 \in R^{l_{offd}\times(l-m-1)}$ that contains the rest of the columns is not needed in the analysis that follows. Applying the elementary row operations described in the proof of Lemma 9 to the first $m+1$ columns of L', we can transform L'_1 to an upper triangular form, with its $(k+1)$th diagonal entry $\lambda_k\Delta x_k$, where $\lambda_k \in Z$ and $k \in \mathcal{A}$. Clearly, any $\lambda_k = 0$ or Δx_k will lead to $M(L') = 0$.

Now let us study some forms of perturbation matrices e for $L' + e$ to lose rank. The case that L' is not full rank is trivial, as $M(L') = 0$. When L' is full rank, $L' + e$ will lose rank if there exists any $\Delta\hat{x}_i = 0$. Assume WLOG that $\Delta\hat{x}_1 = 0$, $e_{x_{j(1)}} = e_{x_{j(2)}} = 0$, for all $j \neq 1, j \in \mathcal{A}$. Let us calculate $\|e\|_F$. Clearly, $e_{x_{1(1)}} = e_{x_{1(2)}} + \Delta x_1$, and

$$\|e\|_F = \sqrt{c_1 e_{x_{1(1)}}^2 + (l_{\text{offd}} - c_1)e_{x_{1(2)}}^2 + \sum_{k=2}^{n}\left\{\left[b_{k(1)}x_{k(1)}^2 + \left(c_1 - b_{k(1)}\right)x_{k(2)}^2\right]e_{x_{1(1)}}^2 + \left[b_{k(2)}x_{k(1)}^2 + \left(l_{\text{offd}} - c_1 - b_{k(2)}\right)x_{k(2)}^2\right]e_{x_{1(2)}}^2\right\} + \cdots}$$

$$= \sqrt{\left(e_{x_{1(2)}} + \Delta x_1\right)^2\left\{c_1 + \sum_{k=2}^{n}\left[b_{k(1)}x_{k(1)}^2 + \left(c_1 - b_{k(1)}\right)x_{k(2)}^2\right] + \cdots\right\} + e_{x_{1(2)}}^2\left\{(l_{\text{offd}} - c_1) + \sum_{k=2}^{n}\left[b_{k(2)}x_{k(1)}^2 + \left(l_{\text{offd}} - c_1 - b_{k(2)}\right)x_{k(2)}^2\right] + \cdots\right\}}$$

$$= \sqrt{e_{x_{1(2)}}^2\left\{l_{\text{offd}} + \sum_{k=2}^{n}\left[\left(b_{k(1)} + b_{k(2)}\right)x_{k(1)}^2 + \left(l_{\text{offd}} - b_{k(1)} - b_{k(2)}\right)x_{k(2)}^2\right] + \cdots\right\} + \left(2e_{x_{1(2)}}\Delta x_1 + \Delta x_1^2\right)\left\{c_1 + \sum_{k=2}^{n}\left[b_{k(1)}x_{k(1)}^2 + \left(c_1 - b_{k(1)}\right)x_{k(2)}^2\right] + \cdots\right\}}$$

where $b_{k(1)}$ and $b_{k(2)}$ are respectively the numbers of $x_{k(1)}x_{1(1)}$ and $x_{k(1)}x_{1(2)}$ in the column that corresponds to $x_k x_1$ in L'. For simplicity, we denote $\left(l_{\text{offd}} + \sum_{k=2}^{m}\left[\left(b_{k(1)} + b_{k(2)}\right)x_{k(1)}^2 + \left(l_{\text{offd}} - b_{k(1)} - b_{k(2)}\right)x_{k(2)}^2\right] + \cdots\right)$ as $B_{1,1}$, and $\left(c_1 + \sum_{k=2}^{m}\left[b_{k(1)}x_{k(1)}^2 + \left(c_1 - b_{k(1)}\right)x_{k(2)}^2\right] + \cdots\right)$ as $B_{1,2}$, $B_{1,1}, B_{1,2} > 0$. Note that $B_{1,1} - B_{1,2} = \left(l_{\text{offd}} - c_1\right) + \sum_{k=2}^{m} b_{k(2)}x_{k(1)}^2 + \left(l_{\text{offd}} - c_1 - b_{k(2)}\right)x_{k(2)}^2 + \cdots \geq 0$. Then,

$$\|e\|_F = \sqrt{e_{x_{1(2)}}^2 B_{1,1} + \left(2e_{x_{1(2)}}\Delta x_1 + \Delta x_1^2\right)B_{1,2}}$$

$$= \sqrt{B_{1,1}\left(e_{x_{1(2)}} + \frac{B_{1,2}}{B_{1,1}}\Delta x_1\right)^2 + \Delta x_1^2\left(B_{1,2} - \frac{B_{1,2}^2}{B_{1,1}}\right)}$$

$$= \sqrt{B_{1,1}\left(e_{x_{1(2)}} + \frac{B_{1,2}}{B_{1,1}}\Delta x_1\right)^2 + \Delta x_1^2\left(\frac{B_{1,2}(B_{1,1} - B_{1,2})}{B_{1,1}}\right)}$$

$$\geq \Delta x_1\sqrt{\frac{B_{1,2}\left(B_{1,1} - B_{1,2}\right)}{B_{1,1}}}$$

The equality holds when $e_{x_{1(1)}} = \frac{B_{1,1}-B_{1,2}}{B_{1,1}}\Delta x_1$ and $e_{x_{1(2)}} = -\frac{B_{1,2}}{B_{1,1}}\Delta x_1$. Note that $\Delta x_1\sqrt{\frac{B_{1,2}\left(B_{1,1}-B_{1,2}\right)}{B_{1,1}}} = \Delta x_1\sqrt{-\frac{1}{B_{1,1}}\left(B_{1,2} - \frac{B_{1,1}}{2}\right)^2 + \frac{B_{1,1}}{4}} \leq \Delta x_1\frac{\sqrt{B_{1,1}}}{2}$, and the equality holds when $B_{1,2} = \frac{B_{1,1}}{2}$, indicating that

$$B_{1,1} = l_{\text{offd}} + \frac{l_{\text{offd}}}{2}\sum_{k=2}^{m}\left(x_{k(1)}^2 + x_{k(2)}^2\right) + \cdots + \frac{l_{\text{offd}}}{2^{\tau-1}}\left[\prod_{k=2}^{m}\left(x_{k(1)}^2 + x_{k(2)}^2\right) + \cdots\right]$$

$$= l_{\text{offd}} + \frac{l_{\text{offd}}}{2}\sum_{j=1}^{\binom{m-1}{1}}\prod_{a\in A_{1,1,j}}\left(x_{a(1)}^2 + x_{a(2)}^2\right) + \cdots + \frac{l_{\text{offd}}}{2^{\tau-1}}\sum_{j=1}^{\binom{m-1}{\tau-1}}\prod_{a\in A_{1,\tau-1,j}}\left(x_{a(1)}^2 + x_{a(2)}^2\right)$$

$$= l_{\text{offd}} + l_{\text{offd}}\sum_{i=1}^{\tau-1}\left\lfloor\frac{1}{2^i}\sum_{j=1}^{\binom{m-1}{i}}\prod_{a\in A_{1,i,j}}\left(x_{a(1)}^2 + x_{a(2)}^2\right)\right\rfloor.$$

Therefore,

$$\Delta x_1\sqrt{\frac{B_{1,2}\left(B_{1,1} - B_{1,2}\right)}{B_{1,1}}} \leq \frac{\sqrt{l_{\text{offd}}}}{2}\Delta x_1\sqrt{1 + \sum_{i=1}^{\tau-1}\left\lfloor\frac{1}{2^i}\sum_{j=1}^{\binom{m-1}{i}}\prod_{a\in A_{1,i,j}}\left(x_{a(1)}^2 + x_{a(2)}^2\right)\right\rfloor}$$

In cases when $\Delta\hat{x}_k = 0$, $k = 1$, we define $B_{k,1}$ and $B_{k,2}$ in a similar way as $B_{1,1}$ and $B_{1,2}$. Then similar results can be obtained, which lead to

$$M(L') \leq \min\left\{\Delta x_k\sqrt{\frac{B_{k,2}\left(B_{k,1} - B_{k,2}\right)}{B_{k,1}}}\right\}$$

$$\leq \frac{\sqrt{l_{\text{offd}}}}{2} \min \left\{ \left\lfloor \Delta x_k \sqrt{1 + \sum_{i=1}^{m-1} \left\lfloor \frac{1}{2^i} \sum_{j=1}^{i} \prod_{a \in A_{k,i,j}} \left(x_{a(1)}^2 + x_{a(2)}^2 \right) \right\rfloor} \right\rfloor, k \in A \right\}$$

$$= M(L'_{\text{offd}})$$

The equality is achieved by an OFFD.

A.12. Proof of Corollary 1

Assume WLOG that the two input parameters x_1 and x_2 in the second and third columns of L' have k out of four combinations of levels, where $k \in \{1, 2\}$. The elementary row operations shown in the proof of Lemma 9 lead to non-zero diagonal entries in the first k rows. After subtraction of all the other $2^{m-y^*} - k$ rows by one of these first k rows, the first three elements in all these other rows become zeros, leading to $\lambda_2 = 0$. $M(L') = 0$ then follows.

A.13. Proof of Corollary 2

Assume WLOG that x_1 and x_2 are the two input parameters which have k out of four combinations of levels, where $k \in \{1, 2, 3\}$. Through switching the column of $x_i x_j$ to the 4th column, the first k rows contain different combinations. After subtraction of the other $2^{m-y^*} - k$ rows by one of the first k rows, the first four elements in these other rows become zeros, leading to $\lambda_3 = 0$. $M(L') = 0$ then follows.

A.14. Proof of Theorem 5

When $m \leq 2$ or $\tau > \lceil \frac{m}{2} \rceil - 1$, no further reduction is possible. The M-PCM simulation set forms a full factorial design, which correctly estimates all effects of input parameters. When $m > 2$, $1 < \tau \leq \lceil \frac{m}{2} \rceil - 1$, and OFFDs are applied to further reduce the size of simulation set. According to the property of OFFDs discussed in Section 2.2 [4,48], a k-factor interaction effect is only confounded with interaction effects of at least $(R - k)$-factors. The conditions $1 \leq k \leq \tau$ and $R \geq 2\tau + 1$ simply lead to $R - k \geq R - \tau \geq \tau + 1 > \tau$. As interaction effects involving more than τ factors do not exist, neither the main effects nor any interaction effect is confounded.

References

[1] B.E. Ankenman, Design of experiments with two-and four level factors, J. Qual. Technol. 31 (4) (1999) 363.
[2] K.E. Atkinson, An Introduction to Numerical Analysis, John Wiley & Sons, 2008.
[3] D. Bigoni, A.P. Engsig-Karup, Y.M. Marzouk, Spectral tensor-train decomposition, SIAM J. Sci. Comput. 38 (4) (2016) A2405–A2439.
[4] G.E. Box, J. Hunter, W.G. Hunter, Statistics for Experimenters, Wiley, New York, 2005.
[5] S. Burge, The Systems Engineering Tool Box, London, UK, 2006, p. 1.
[6] M. Cameron, Gaussian quadrature, 2018, http://www2.math.umd.edu/ ~mariakc/teaching/gaussian.pdf. (Accessed on June 8, 2018).
[7] M. Chen, P. Wang, Multi-level factorial designs with minimum numbers of level changes, Taylor & Francis, 2001.
[8] H. Cheng, A. Sandu, Efficient uncertainty quantification with the polynomial chaos method for stiff systems, Math. Comput. Simulation 79 (11) (2009) 3278–3295.
[9] W. Conner, M. Zelen, L. Deming, Fractional factorial experimental designs for factors at two levels, Natl. Bur. Stand. Appl. Math. Ser. 48 (1956).
[10] P.R. Conrad, Y.M. Marzouk, Adaptive smolyak pseudospectral approximations, SIAM J. Sci. Comput. 35 (6) (2013) A2643–A2670.
[11] P.G. Constantine, M.S. Eldred, E.T. Phipps, Sparse pseudospectral approximation method, Comput. Methods Appl. Mech. Engrg. 229 (2012) 1–12.
[12] M. Dahleh, M.A. Dahleh, G. Verghese, Lectures on dynamic systems and control, A+ A 4 (100) (2004) 1–100.
[13] J.W. Demmel, On condition numbers and the distance to the nearest ill-posed problem, Numer. Math. 51 (3) (1987) 251–289.
[14] P. Diaz, A. Doostan, J. Hampton, Sparse polynomial chaos expansions via compressed sensing and d-optimal design, Comput. Methods Appl. Mech. Engrg. 336 (2018) 640–666.
[15] A. Doostan, A. Validi, G. Iaccarino, Non-intrusive low-rank separated approximation of high-dimensional stochastic models, Comput. Methods Appl. Mech. Engrg. 263 (2013) 42–55.

[16] J. Filliben, E. Simiu, Tall building response parameters: sensitivity study based on orthogonal factorial experiment design technique, J. Struct. Eng. 136 (2) (2010) 160–164.
[17] A. Gil, J. Segura, N.M. Temme, Numerical Methods for Special Functions, SIAM, 2007.
[18] M.B. Giles, Multilevel monte carlo methods, Acta Numer. 24 (2015) 259–328.
[19] P.W. Glynn, D.L. Iglehart, Importance sampling for stochastic simulations, Manage. Sci. 35 (11) (1989) 1367–1392.
[20] A.A. Gorodetsky, Continuous low-rank tensor decompositions, with applications to stochastic optimal control and data assimilation, Ph.D. Thesis, Massachusetts Institute of Technology, 2017.
[21] A.A. Gorodetsky, S. Karaman, Y.M. Marzouk, Function-train: A continuous analogue of the tensor-train decomposition, 2015, arXiv preprint arXiv:1510.09088.
[22] R.F. Gunst, R.L. Mason, Fractional factorial design, Wiley Interdiscip. Rev. Comput. Stat. 1 (2) (2009) 234–244.
[23] M. Gunzburger, C.G. Webster, G. Zhang, Sparse collocation methods for stochastic interpolation and quadrature, in: Handbook of Uncertainty Quantification, Springer, 2016, pp. 1–46.
[24] T. Hachisuka, W. Jarosz, R.P. Weistroffer, K. Dale, G. Humphreys, M. Zwicker, H.W. Jensen, Multidimensional adaptive sampling and reconstruction for ray tracing, in: ACM Transactions on Graphics, TOG, vol. 27, ACM, 2008, p. 33.
[25] J. Hampton, A. Doostan, Coherence motivated sampling and convergence analysis of least squares polynomial chaos regression, Comput. Methods Appl. Mech. Engrg. 290 (2015) 73–97.
[26] J. Hampton, A. Doostan, Compressive sampling of polynomial chaos expansions: Convergence analysis and sampling strategies, J. Comput. Phys. 280 (2015) 363–386.
[27] J. Hampton, A. Doostan, Basis adaptive sample efficient polynomial chaos (base-pc), J. Comput. Phys. 371 (2018) 20–49.
[28] N. Haritos, Monte carlo simulation of ocean beacon response to environmental loading, Math. Comput. Simulation 30 (1–2) (1988) 87–92.
[29] F. Heiss, V. Winschel, Likelihood approximation by numerical integration on sparse grids, J. Econometrics 144 (1) (2008) 62–80.
[30] J.S. Hesthaven, B. Stamm, S. Zhang, Efficient greedy algorithms for high-dimensional parameter spaces with applications to empirical interpolation and reduced basis methods, ESAIM Math. Model. Numer. Anal. 48 (01) (2014) 259–283.
[31] J.R. Hockenberry, B.C. Lesieutre, Evaluation of uncertainty in dynamic simulations of power system models: The probabilistic collocation method, IEEE Trans. Power Syst. 19 (3) (2004) 1483–1491.
[32] S.S. Isukapalli, Uncertainty analysis of transport-transformation models, Ph.D. Thesis, Rutgers, The State University of New Jersey, 1999.
[33] T. Kato, Perturbation Theory for Linear Operators, vol. 132, Springer, 1995.
[34] A.I. Khuri, S. Mukhopadhyay, Response surface methodology, Wiley Interdiscip. Rev. Comput. Stat. 2 (2) (2010) 128–149.
[35] P. Knupp, Matrix norms and the condition number, in: Proceedings of 8th International Meshing Roundtable, Citeseer, South Lake Tahoe, CA, 1999, pp. 13–22.
[36] J. Kober, J.A. Bagnell, J. Peters, Reinforcement learning in robotics: A survey, Int. J. Robot. Res. 32 (11) (2013) 1238–1274.
[37] S.H. Lee, W. Chen, A comparative study of uncertainty propagation methods for black-box-type problems, Struct. Multidiscip. Optim. 37 (3) (2009) 239–253.
[38] Y. Lei, R. Kacker, D. Kuhn, V. Okun, J. Lawrence, IPOG/IPOD: Efficient test generation for multi-way software testing, J. Softw. Test. Verif. Reliab. 18 (3) (2008) 125–148.
[39] G. Li, S. Wang, H. Rabitz, High dimensional model representations (HDMR): Concepts and applications, in: Proceedings of the Institute of Mathematics and Its Applicatons Workshop on Atmospheric Modeling, Citeseer, 2000, pp. 15–19.
[40] Y. Matsuno, T. Tsuchiya, J. Wei, I. Hwang, N. Matayoshi, Stochastic optimal control for aircraft conflict resolution under wind uncertainty, Aerosp. Sci. Technol. 43 (2015) 77–88.
[41] M.D. McKay, R.J. Beckman, W.J. Conover, A comparison of three methods for selecting values of input variables in the analysis of output from a computer code, Technometrics 42 (1) (2000) 55–61.
[42] K.L. Mills, J.J. Filliben, Comparison of two dimension-reduction methods for network simulation models, J. Res. Natl. Inst. Stand. Technol. 116 (5) (2011) 771.
[43] D.C. Montgomery, Design and Analysis of Experiments, vol. 7, Wiley, New York, 1997.
[44] F. Nobile, R. Tempone, C.G. Webster, An anisotropic sparse grid stochastic collocation method for partial differential equations with random input data, SIAM J. Numer. Anal. 46 (5) (2008) 2411–2442.
[45] I.V. Oseledets, Tensor-train decomposition, SIAM J. Sci. Comput. 33 (5) (2011) 2295–2317.
[46] T.N.L. Patterson, The optimum addition of points to quadrature formulae, Math. Comp. 22 (104) (1968) 847–856.
[47] J. Peng, J. Hampton, A. Doostan, On polynomial chaos expansion via gradient–enhanced 1-minimization, J. Comput. Phys. 310 (2016) 440–458.
[48] J. Prinz, P. Tobias, W.F. Guthrie, B. Hembree, M.C. Croarkin, J.J. Filliben, N.A. Heckert, NIST/SEMATECH e-handbook of statistical methods, in: NIST Handbook 151, 2013, Chap V, Section 5, Subsection 9. http://www.itl.nist.gov/div898/handbook/.
[49] M. Robert, A Comprehensive Guide to Factorial Two-Level Experimentation, Springer, 2009.
[50] A. Saltelli, K. Chan, E. Scott, et al., Sensitivity Analysis vol. 1, Wiley, New York, 2000.
[51] A. Saltelli, M. Ratto, T. Andres, F. Campolongo, J. Cariboni, D. Gatelli, M. Saisana, S. Tarantola, Global Sensitivity Analysis: The Primer, John Wiley & Sons, 2008.
[52] D.V. Savostyanov, Quasioptimality of maximum-volume cross interpolation of tensors, Linear Algebra Appl. 458 (2014) 217–244.
[53] T. Shirakura, Fractional factorial designs of two and three levels, Discrete Math. 116 (1–3) (1993) 99–135.
[54] I.M. Sobol, Sensitivity estimates for nonlinear mathematical models, Math. Modelling Comput. Exp. 1 (4) (1993) 407–414.
[55] J.A. Sokolowski, C.M. Banks, Modeling and Simulation Fundamentals: Theoretical Underpinnings and Practical Domains, John Wiley & Sons, 2010.
[56] L.E. Spence, A.J. Insel, S.H. Friedberg, Elementary Linear Algebra, Pearson/Prentice Hall, 2008.
[57] J.M. Steele, The Cauchy-Schwarz Master Class: An Introduction to the Art of Mathematical Inequalities, Cambridge University Press, 2004.

[58] M.A. Tatang, Direct incorporation of uncertainty in chemical and environmental engineering systems, Ph.D. Thesis, Massachusetts Institute of Technology, 1995, (Chapter 7.4).
[59] M.A. Tatang, W. Pan, R.G. Prinn, G.J. McRae, An efficient method for parametric uncertainty analysis of numerical geophysical models, J. Geophys. Res. Atmos. (1984–2012) 102 (D18) (1997) 21925–21932.
[60] L. Trutna, et al., Summary tables of useful fractional factorial designs, 2013, http://www.itl.nist.gov/div898/handbook/pri/section3/pri3347.htm. (Last Modified Oct. 30, 2013).
[61] Y. Wan, S. Roy, B. Lesieutre, Uncertainty evaluation through mapping identification in intensive dynamic simulations, IEEE Trans. Syst. Man Cybern. A 40 (5) (2010) 1094–1104.
[62] S. Wang, G. Huang, W. Huang, Y. Fan, Z. Li, A fractional factorial probabilistic collocation method for uncertainty propagation of hydrologic model parameters in a reduced dimensional space, J. Hydrol. 529 (2015) 1129–1146.
[63] D. Wei, Z. Cui, J. Chen, Uncertainty quantification using polynomial chaos expansion with points of monomial cubature rules, Comput. Struct. 86 (23) (2008) 2102–2108.
[64] S.J. Wetzel, C.M. Guttman, K.M. Flynn, J.J. Filliben, Significant parameters in the optimization of maldi-tof-ms for synthetic polymers, J. Am. Soc. Mass Spectrom. 17 (2) (2006) 246–252.
[65] J. Xie, Y. Wan, Scalable multidimensional uncertainty evaluation approach to strategic air traffic flow management, in: AIAA modeling and simulation technologies conference, 2015, p. 2492.
[66] J. Xie, Y. Wan, F. Lewis, Strategic air traffic flow management under uncertainties using scalable sampling-based dynamic programming and q-learning approaches, in: Control Conference, ASCC, 2017 11th Asian, IEEE, 2017, pp. 1116–1121.
[67] J. Xie, Y. Wan, K. Mills, J.J. Filliben, F. Lewis, A scalable sampling method to high-dimensional uncertainties for optimal and reinforcement learning-based controls, IEEE Control Syst. Lett. 1 (1) (2017) 98–103.
[68] J. Xie, Y. Wan, Y. Zhou, K. Mills, J.J. Filliben, Y. Lei, Effective and scalable uncertainty evaluation for large-scale complex system applications, in: Winter Simulation Conference, WSC, IEEE, 2014, pp. 733–744, Savannah, GA.
[69] D. Xiu, G.E. Karniadakis, The Wiener–Askey polynomial chaos for stochastic differential equations, SIAM J. Sci. Comput. 24 (2) (2002) 619–644.
[70] Y. Zhou, Y. Wan, S. Roy, C. Taylor, C. Wanke, Multivariate probabilistic collocation method for effective uncertainty evaluation with application to air traffic management, in: IEEE Conference on Decision and Control, Maui, HI, 2012.
[71] Y. Zhou, Y. Wan, S. Roy, C. Taylor, C. Wanke, D. Ramamurthy, J. Xie, Multivariate probabilistic collocation method for effective uncertainty evaluation with application to air traffic flow management, IEEE Trans. Syst. Man Cybern. A 44 (10) (2014) 1347–1363.
[72] Y. Zhou, Y. Wan, C. Taylor, S. Roy, C. Wanke, Performance evaluation and optimal decision-making for strategic air traffic management under weather uncertainty, Proceedings of AIAA Infotech Aerospace Conference, Garden Grove, CA, 2012.
[73] Y. Zhou, Y. Wan, C. Wanke, C. Taylor, S. Roy, A probabilistic collocation method-based approach for optimal strategic air traffic flow management under weather uncertainties, Proceedings of AIAA ATIO Conference, Los Angeles, CA, 2013.